Education-Based Cutoffs for Cognitive Screening of þÿ A l z h e i m e r s D i s e a s e

Alenius, Minna M

2022

Alenius , M M , Ngandu , T , Koskinen , S , Hallikainen , I , Hänninen , T , Karrasch , M , Kivipelto , M , Raivio , M , Laakkonen , M-L , Kruger , J , Suhonen , N M & Hokkanen , L þÿ 2 0 2 2 , ' E d u c a t i o n - B a s e d C u t o f f s f o r C o g n i t i v e S c r e e n i n g o f A l z h e i m e r
Dementia and Geriatric Cognitive Disorders , vol. 51 , no. 1 , pp. 42-55 . https://doi.org/10.1159/000521982

http://hdl.handle.net/10138/352932
https://doi.org/10.1159/000521982

unspecified
acceptedVersion

Research Article

Education-based cut-offs for cognitive screening of Alzheimer's disease

Minna Alenius[a,b], Tiia Ngandu[b,c], Sanna Koskinen[a], Ilona Hallikainen[d], Tuomo Hänninen[d,e], Mira Karrasch[f], Miia Kivipelto[b,c,g,h], Minna M Raivio[i,j], Marja-Liisa Laakkonen[i,j,k], Johanna Krüger[l,m], Noora-Maria Suhonen[n], Laura Hokkanen[a]

[a] Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland
[b] Population Health Unit, Finnish Institute for Health and Welfare, Helsinki, Finland
[c] Division of Clinical Geriatrics, Center for Alzheimer Research, NVS, Karolinska Institutet, Stockholm, Sweden
[d] Department of Neurology, Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland
[e] Neurology of NeuroCenter, Kuopio University Hospital, Kuopio, Finland
[f] Department of Psychology, Åbo Akademi University, Turku, Finland
[g] Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland
[h] Ageing Epidemiology Research Unit, School of Public Health, Imperial College London, London, United Kingdom
[i] Department of General Practice, University of Helsinki, Helsinki, Finland
[j] Unit of Primary Health Care, Helsinki University Hospital, Helsinki, Finland
[k] Geriatric Clinic, Department of Social Services and Health Care, Laakso Hospital, Helsinki, Finland
[l] Research Unit of Clinical Neuroscience, Neurology, University of Oulu, Oulu, Finland
[m] MRC, Oulu University Hospital, Oulu, Finland
[n] Department of Neurology, Oulu University Hospital, Oulu, Finland

Short title: Education-based cut-offs for cognitive screening

Corresponding author:

Minna Alenius

Department of Psychology and Logopedics

PO. Box 9

FI-00014 University of Helsinki, Finland

Tel: +358-407335331

E-mail: minna.alenius@helsinki.fi

Number of tables: 8

Number of figures: 1

Number of supplemental materials: 1 table, 1 figure

Word count: Total 3574 words; without brackets indicating places for tables and figures 3550 words.

Keywords: Cognitive performance, Alzheimer’s disease, CERAD, cut-off, education

Abstract

Introduction: The educational background and size of the elderly population are undergoing significant changes in Finland during the 2020s. A similar process is likely to occur also in several European countries. For cognitive screening of early Alzheimer's disease (AD) using outdated norms and cut-off scores may negatively affect clinical accuracy. The aim of the present study was to examine the effects of education, age, and gender on the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuropsychological battery (CERAD-nb) in a large register-based, clinical sample of patients with mild AD and non-demented at-risk persons from the general population (controls) and to examine whether corrected cut-off scores would increase the accuracy of differentiation between the two groups. **Methods:** CERAD-nb scores were obtained from AD patients (n = 389, 58% women, mean age 74.0 years) and from controls (n = 1980, 52% women, mean age 68.5 years). The differences in CERAD-nb performance were evaluated by univariate GLM. Differentiation between the two groups was evaluated using a receiver operating characteristics (ROC) curve, where a larger area under the ROC curve (AUC) represents better discrimination, Youden's J was calculated for the overall performance and accuracy of each of the measures. **Results:** Of the demographic factors, education was the strongest predictor of CERAD-nb performance, explaining more variation than age or gender in both the AD patients as well as in the controls. Education corrected cut-off scores had better diagnostic accuracy in discriminating between the AD patients and controls than existing uncorrected scores. The highest level of discrimination between the two groups overall were found for the two CERAD-nb total scores. **Conclusions**: Education-corrected cut-off scores were superior to uncorrected scores in differentiating between controls and AD patients especially for the highest level of education and should therefore be used in clinical cognitive screening; also, as the proportion of the educated elderly is increasing substantially during 2020s. Our results also indicate that total scores of the CERAD-nb are better at discriminating AD patients from controls than any single subtest score. A digital tool for calculating the total scores and comparing education-based cut-offs would increase the efficiency and usability of the test.

Introduction

Cognitive performance declines with age, with an increasingly steeper decline in old age [1–6], across diverse ethnocultural groups and geographical regions [7]. For people aged 60 years and over, the prevalence of dementia increases exponentially with age, Alzheimer's disease (AD) being the major cause of dementia [8]. In addition to the effects of age on cognitive performance, there may also be cohort-related effects related to historical influences such as possibilities for education, assumed gender roles, cultural factors, and socioeconomic status [9]. Throughout the European countries, substantial differences exist in cognitive performance among the elderly [10].

Early cognitive predictors of dementia differ according to educational level [11]. Higher education level is associated with better cognitive performance in several cognitive measures[1,3–6,12]. Education is reported to have a protective effect on cognitive function by improving cognitive reserve (CR) [12,13]. Individuals with higher education have a lower prevalence or incidence of dementia

and, if dementia is diagnosed, higher cognitive performance at initial assessment [13]. Recently, new guidelines and consensus criteria for CR and its evaluation have been published [14], introducing a recommendation to include all three components, i.e., structural status of the brain, clinical or cognitive performance outcomes, and a measure of reserve, in the CR evaluation.

Cognitive performance and CR may also differ according to gender. Gender differences may vary systematically across birth cohorts and regions and can be associated with changes in living conditions and educational opportunities [15]. Some studies have reported that execution in visuospatial tasks is mainly dependent on gender; the superiority of healthy men compared with women in tasks of visual memory [12,16], object naming, and constructional praxis [17] seems to remain stable even when people develop AD. One study also reports that this indicates larger CR in men [17]. Additionally, some studies report that cognitively healthy women outperform men in tests of verbal memory and verbal fluency [1,2,5], but these results may not be stable in AD [17].

The population is aging worldwide. In Finland, the total proportion of those aged 65 years and over (elderly) is growing during the 2020s and the proportion of those aged 75 years and over (group ≥75) is growing even faster [18,19]. At the same time, the educational background of the elderly and the group ≥75 is also changing [20]; supplemental materials describe both phenomena. The increase and structural transformation of the elderly will challenge the Finnish cognitive dementia screening processes, which originally were implemented for a smaller number of elderly citizens in 1999 [18,21].

[Supplemental FIGURE 1]

[Supplemental TABLE 1]

Educational opportunities for Finnish cohorts could be classified into the following three groups [22]: 1. the war generation, born 1935 or before, with scant education; 2. the generation of structural change, born 1936–1955, with growing educational opportunities; and 3. the welfare generation, born after 1955, with many educational choices. The war generation followed the 1921 educational legislation with a basic educational right and duty for everyone aged between 7 and 12 years [23]. The structural change generation had a school system in which, after reaching 12 years of age, students had the possibility to either continue for two more years in secondary school or to switch over to a two-year civic school [23]. For this generation, schools granting qualifications in manual industry occupations were also available [23]. The welfare generation has followed a comprehensive

school system, including 9 years of compulsory education up to the age of 16, for cohorts born from the early to mid-1960s onwards [23].

Originally, the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) was established in the 1980s to standardize and validate measures for the assessment of AD, and the neuropsychological battery (CERAD-nb) was introduced as a tool to measure early cognitive impairment in AD [24]. Later, an index of overall level of cognitive functioning on the CERAD-nb (Chandler Total) was introduced [25], and another index (Seo Total) expanded the Chandler Total by adding constructional praxis recall to it [26]. Both scores increase the utility of the CERAD-nb [27–30]. In Finland, the CERAD-nb has been in use for more than 20 years [21]. During 2006-2008 the national best practices in the treatment of progressive memory disorders were upgraded and new recommendations were published [31]. Focus was put on early diagnosis, treatment, and care, with the focus shifting from special care to multi-professional memory teams in primary care. For cognitive screening the CERAD-nb replaced MMSE and was changed from an assessment tool to a cognitive screening tool used mainly in primary care by memory nurses. The national CERAD-nb reference values were introduced in 2007 [33] and the cut-off scores in 2010 [34,35]. These 2010 cut-offs scores are uniform and thus education, age or gender are not corrected for.

As can be seen in the supplemental materials, the proportion of the educated elderly is increasing considerably in Finland during the 2020s – the proportion with the lowest educational level has decreased from 55% in the cohort born <1945 to 13% in the cohort born 1965-1969 [20]. In the highest educational group, the proportion of women exceeds the proportion of men with an increasing proportion overall from the cohort born 1950-1954 onwards [20]. At the end of the 2020s the group aged ≥75 years will be roughly 1.5 times larger than in 2019[18,19]. A similar process is likely to occur in the near future in several European countries [36]. The overarching aim in the

present study was to evaluate the risk of outdated cut-off scores and need for renewing education-, age-, gender-based cut-offs in early diagnosis of AD. The specific research questions were as follows:

1. What is the effect of education, age, and gender on CERAD-nb in this population?
2. How does cognitive performance differ in subgroups by education, age, and gender?
3. Do the results suggest clinical renewal of the Finnish CERAD-nb cut-offs for improved discriminating power?

Materials and Methods

Study population

Patients with their first recorded AD diagnosis during the years 2010–2013 were identified from primary and secondary care treatment notifications of the Finnish national health registers (NHR). Patients aged 60–80 years at the time of AD diagnosis in the Helsinki, Turku, Kuopio, Seinäjoki, and Oulu areas were targeted. Patients were excluded if before their AD diagnosis they had a diagnosis of dementia other than AD, other major neurological or psychiatric or developmental disorders impacting cognition, or diagnosis of serious mood disorders or alcoholism or other transitory psychiatric or neurological disorders affecting cognition during the previous 24 months. Additionally, sampling from a private sector provider (PSP) in Helsinki was added because roughly 7% of the NHR-based sample turned out to be initially diagnosed in the private sector, mainly in the Helsinki area. The sampling, in terms of age, gender, and geographic location, focused on patients comparable to the participants of the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) [37], which formed the non-demented comparison group (controls). After AD diagnosis was identified from the NHR registries, further information was requested from local health records (LHR). This study was approved by the Finnish Institute for Health and Welfare (THL), decision THL/1649/6.02.00/2018; also, local approvals were obtained as needed. The final AD cohort of cognitive performance evaluations with the CERAD-nb consisted of 389 patients. Figure 1 summarizes the sampling.

The controls comprised persons who participated during 2009–2011 in the screening visit of the FINGER study [34]. They were aged 60–77 years, not demented, and had a Cardiovascular Risk Factors, Aging and Incidence of Dementia (CAIDE) risk score of 6 points or higher, indicating the presence of some modifiable risk factors. Characteristics of the FINGER participants (ClinicalTrials.gov identifier: NCT01041989) have been described earlier [37]. The controls comprised 1980 persons after exclusions due to incomplete information, conditions affecting cognitive performance, and

geographical area not included. Both control and AD cohorts included patients from urban and rural areas from the south, east, north, central, and west parts of Finland.

Evaluating eligibility of AD patients

The inclusion criteria for cognitive performance evaluations were mild state of AD, and the availability of Mini Mental State Examination (MMSE) and the majority of the CERAD-nb measures assessed less than 12 months prior to AD diagnosis. Subtest components were not obligatory.

AD was diagnosed according to Finnish Current Care Guidelines for Memory Diseases [38], which are based on recommendations of the NINCDS-ARDRA [39] and DSM-IV [40] criteria. AD state was evaluated with the combination of MMSE [41] and either the Global Deterioration Scale (GDS) [42] or the Clinical Dementia Rating (CDR) scale [43] or written descriptions of daily activities capability. For cases with only MMSE available (n=115), a score of 20–30 was used to define mild AD, 10–19 moderate AD, and <10 severe AD.

[FIGURE 1]

Classification of education

Educational attainment for the controls was reported both in years and with the highest degree. For the AD group, the highest degree was available for 294 patients, for some (n=75) only occupational information was available, and for the minority (n=20) no information on education was available. Thus, the classification of education was based roughly on school years; estimated or given primary school 7 years or less (lowest level); approximately 8–10 years of education, for example, studies in middle/technical/trade/vocational schools or matriculation examination without any other degree (middle level); or more than 10 years with a degree from college or university (highest level). For AD patients with only occupational information, the educational level was estimated based on occupation; the majority belonged to the lowest level occupations like cleaner, bus driver, and storeman.

The CERAD-nb

The CERAD-nb [24] is a relatively brief (i.e., 20–30 minutes) assessment tool, including the following subtests: Verbal Fluency (animal category), 15-item Boston Naming Test, MMSE, Word List Memory (sum of 3 trials of 10 words each), Constructional Praxis Copy (including circle, diamond, overlapping rectangles, cube), Word List Recall and Word List Recognition, and Constructional Praxis Recall. The Finnish version of the CERAD-nb also includes the Clock Drawing subtest [33]. Additional variables were also calculated: Word List Savings is the proportion (max 100%) of words retrieved in delayed recall compared with the last learning trial; Word List Recognition is both correct yes and no

responses scored together and then calculated as a proportion of total 20; Verbal Memory total recall is the sum of word list delayed recall score and word list recognition score; Constructional Praxis Savings is the proportion (max 100%) of the four figures retrieved from memory compared with the copy. Chandler Total [25] sums up Verbal Fluency (max 24), 15-item Boston Naming Test, Word List Memory, Constructional Praxis Copy, Word List Recall, and Word List Recognition correct yes - false yes. Seo Total [26] sums up Chandler Total with Constructional Praxis Recall.

Statistical analyses

To address differences between demographic variables between the AD group and the controls, differences regarding gender were analyzed by Pearson's Chi-square test; education, age group, cohort, and generation differences were analyzed by the Kruskall-Wallis test. The differences in CERAD-nb performance between controls and the AD group were evaluated by univariate general linear models (GLM) controlling for education level, age group, and gender, along with interactions between these variables. To address 1st study question, evaluation of the effects involved the univariate GLM with both separate Eta Squared η^2 for each independent factor as well as Partial Eta Squared η_p^2 for the combinations. To address 2nd study question, the differences in CERAD-nb performance between education levels, age groups, and genders were evaluated by univariate GLM post-hoc test. Performance of pairwise comparisons used Tukey correction. To address the 3rd study question, the ability of each of the subtests to discriminate between controls and the AD group was evaluated using a receiver operating characteristics (ROC) curve, where a larger area under the ROC curve (AUC) represents better discrimination Youden's J was calculated for the overall performance and accuracy of each of the measures. A p-value <0.05 was considered statistically significant, and for univariate GLM corrected p-values were used throughout the study. Statistical analyses were performed using SPSS 26.0 and 27.0 (IBM Corp. Released 2019).

Results

The majority in both patients and controls were women. The mean age for the AD group was 74.0 years (SD 4.7) and for the controls 68.5 years (SD 4.8). Table 1 presents the demographic data for the controls and for the AD group.

[TABLE 1]

As expected, the controls performed better (p<.05) than the AD group on all cognitive measures except for Constructional Praxis Copy. Effect sizes of the difference between groups (AD vs. control),

as shown with Partial Eta Squared column in Table 2, was largest for the total scores, very low for the 15-item Boston Naming Test and Clock Drawing and extremely low for Constructional Praxis Copy.

[TABLE 2]

Effect of demographic variables on cognitive performance

For both the AD sample and the controls, the total effect sizes of the combined model including education, age and gender were largest for both total scores and the 15-item Boston Naming Test. Education had the largest, age a bit smaller, and gender the smallest effect size. Proportion of variance as Eta Squared η^2 ($p < .05$) for each separate model of Education, Age and Gender; and as Partial Eta Squared η_p^2 for combined model of Education, Age, and Gender are presented for controls in Table 3a and for patients with AD diagnosis in Table 3b.

[TABLE 3a]

[TABLE 3b]

The performance differences within education levels, the highest level performing better than the lowest, were found in both samples for all measures; in the controls the highest level also outperformed the middle level for all other measures than the Clock Drawing. The performance differences within age groups, with the younger outperforming the older, were found for all measures in the controls but in the AD group only for half of the measures the performance of the oldest was lower than the other groups. For both groups men performed better for the 15-item Boston Naming Test, Constructional Praxis Recall and Savings (%); in the controls women performed better than men for seven measures, out of which for the AD group five disappeared and two changed for the favor of men. Table 2 provides detailed descriptive statistics for cognitive performance by totals and Table 4a by education, Table 4b by age, and Table 4c by gender.

[TABLE 4a]

[TABLE 4b]

[TABLE 4c]

Discrimination between the groups and the effect of cut-off-values on discrimination accuracy

The most efficient measures for discriminating between the AD group (n=389) and the controls (n=1980), based on AUC-values, were Seo Total (0.954) with a sensitivity of 0.913 and a specificity of 0.844, and Chandler Total (0.951) with sensitivity of 0.906 and specificity of 0.849. Other measures with high discrimination (AUC $\approx$0.900), were Word List Recall, Word List Memory, Verbal Memory Total Recall, Verbal Fluency, and MMSE. Measures with low discrimination (AUC < 0.700), were Constructional Praxis Copy and Clock Drawing. When comparing the 2010 cut-offs and the

uncorrected best Youden-based cut-offs, Word List Memory (17) and Recall (5) were unchanged. Higher rates of false negatives (sensitivity – 1) were found for several measures when using the 2010 cut-offs in the present sample, while the rate of false positives (specificity – 1) was unaffected except for Word List Savings. As education had the largest effect size and highest-education outperformed others, we examined the rates of false negatives and false positives for the participants with the highest education level using uncorrected and education-corrected cut-off scores. The rates of both false negatives and false positives decreased in a majority of tests when using education-based cut-off scores. Table 5 summarizes for the whole sample and for groups stratified by education, age, and gender, the potential cut-offs according to the best Youden's J and potential false negative and positive values if 2010 cut-offs were retained or if education-corrected cut-offs were not used for subjects with the highest-education level.

[TABLE 5]

Discussion/Conclusion

As the educational level of elderly is rapidly rising in Finland and presumably in other European countries as well, it is of importance to examine the potential effects of this demographic change on the cognitive tools used for dementia screening. Our aim was to examine the effect of demographic factors (education, age, gender) on CERAD-nb performances in a large sample of patients with mild AD and controls. We also wanted to examine the potential impact of using different cut-off scores (corrected vs uncorrected for demographic factors) for differentiating between AD patients and controls.

Before addressing the results pertaining to the main aims of our study, it was of clinical significance to note that the two total scores of the CERAD-nb (Chandler Total and Seo Total) had the highest accuracies in discriminating between the AD patients and controls. The AUC-values were higher for these scores than any of the CERAD subtests. Our results support the notion raised by previous studies that the total scores would be a valuable addition in the clinical screening use of the CERAD-nb [25–30].

Demographic effects and cognitive performance

Education was the demographic factor that had the largest effect on cognitive performance, it explained more variation than age and gender. Those with highest education systematically outperformed those with lowest education, this was seen both in the controls and in the AD patients. This result is in line with previous studies in the healthy elderly [2–6,11,12].

Some gender-effects were also observed, with differences between men and women going in both directions. Gender had the largest independent effect size for the 15-item Boston Naming Test in both controls and the AD group, with men performing better. Men have also previously been found to outperform women on object naming [17]. On the other hand, in the control group, women outperformed men on verbal memory measures, as also reported elsewhere [1,2,5,7]. The same did not hold true for AD patients, however, which is in line with the study of Tensil et al. [17]. It is noteworthy that gender-differences in verbal memory functions might also be due to differing educational opportunities of generations and individuals, especially for girls [9,15,20,23]. Several interactions between education, age, and gender were present and could potentially be explained by the profound educational changes [20] shown in the Supplemental materials.

Accuracy of cut-off scores

Our results show that using the 2010 cut off-scores for the Finnish CERAD-nb result in high rates of false negatives (ranging between ~14% to 46% for different subtests). This is highly problematic as sensitivity is important when screening for early stages of AD. Additionally, education-corrected cut-off scores were superior to uncorrected scores in differentiating between controls and AD patients especially for the highest level of education and should therefore be used in clinical cognitive screening; also, as the proportion of the educated elderly is increasing substantially during 2020s. These results are in line with a previous study showing that controlling for education increases the sensitivity of the CERAD for detecting mild cognitive impairment and dementia in patients with Parkinson's disease [44].

For revision of the CERAD-nb cut-offs, analysis of the study population or comparison with earlier research alone is insufficient. The structure and cohort differences of the target population for whom the new cut-offs are being set need to be analyzed as well. As shown in the Supplemental materials, after data collection of the 2010 cut-off study [33–35], the structure of the Finnish elderly population has changed substantially [18] and is continuing to change [19,20].

Additionally, from a practical viewpoint, as manual calculations are laborious and time consuming, the usability of the CERAD-nb for screening purposes could be improved by creating digital tools. The digital tool could be used for calculations as well as automatic flagging of values falling below individually determined cut-off values (e.g., corrected for education).

Strengths and limitations

The main strength of our study is the wide-ranging sampling of patients, covering the national, local, and private sector providers and both urban and rural areas. Our real-world AD population consisted

of all kinds of clinical patients compared with specialized clinic-based studies or research cohorts, which may include biases of diagnostic groups or educational or health conditions.

This study has some limitations. First, some patients with a moderate state of AD (but MMSE>=20) might have been included, as information of GDS/CDR/written capability for daily activities was not available for 29% of the included AD group. Second, the major cohorts/generations of the AD group (born 1930–1939, 74%) and controls (born 1940–1949, 62%) differed, which might have impacted the differences in effect sizes and performance. Third, the controls were higher educated than the AD group. Fourth, the number of AD patients in the age group 60–69 years was small, and the results for this group need to be interpreted with caution.

Conclusion

Education had a statistically and clinically significant effect on CERAD-nb test performances. Using education-corrected cut-off-scores for differentiating between patients with mild AD and controls increased the accuracy of discrimination. It also decreased the rate of both false positives and false negatives in those with the highest level of education. It is also important to note that the two total scores (Chandler Total and Seo Total) were the most efficient measures in discriminating between the AD group from controls. A digital tool for calculating the total scores and comparing education-based cut-offs would increase the efficiency and usability of the CERAD-nb. As the educational background and size of the target population for whom the cut-offs are being set are significantly changing in Finland during the 2020s, updating the cut-offs is warranted.

Statements

Acknowledgments

We thank all LHR and PSP service providers and the persons involved in data collection of the AD group; special acknowledgement is extended to Pirjo Iso-Junno/ Kuopio primary care for excellent support with the local system, and to Markku Nissilä/ Terveystalo for enabling data gathering during the holiday season in July 2020. We thank the members and participants of the FINGER study for their cooperation and for data collection and management.

Statement of Ethics

This register-based study was approved by the Finnish Institute for Health and Welfare (THL), decision THL/1649/6.02.00/2018. As a register-based study, no informed consent was separately obtained. For the controls, a statement of ethics concerning the FINGER participants (with Clinical Trials identifier: NCT01041989) has been provided elsewhere [34].

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

Minna Alenius was supported during 2020 by the Finnish Brain Foundation and during 2021 by the Finnish Alzheimer's Disease Research Society.

FINGER data collection was funded by the Academy of Finland, Finnish Social Insurance Institution, Juho Vainio Foundation, Yrjö Jahnsson Foundation and Alzheimer’s Research and Prevention Foundation.

Author Contributions

TN and MA prepared the study design and first sampling procedure. MA was responsible for data acquisition, data management, and statistical analyses. TN, SK, LH, and MA prepared the conceptualization for key content. MA wrote the first manuscript draft for review and comments of TN, SK, and LH. Other participants reviewed the text and provided input to the interpretation of the results. Additionally, TH, MKa, and LH participated in the pioneer studies of Finnish cognitive screening; TH and LH attended to the studies of 2010-cut-offs – they all proportioned their expertise for the text. All authors revised several manuscript drafts. Finally, all authors read and approved the final manuscript.

Data Availability Statement

Data can be made available only for those fulfilling the requirements for viewing confidential data as required by Finnish law and the Finnish Institute for Health and Welfare. Moreover, the purpose of the research must be in alignment with the informed consent provided for this study and/or the FINGER study (controls), with Finnish law and regulations at the Finnish Institute for Health and Welfare. Requests are to be submitted to the Finnish Institute for Health and Welfare: kirjaamo@thl.fi.

References

1 Cetty L, Peh CX, Abdin E, Vaingankar JA, Shahwan S, Picco L, et al. Normative data for the 10/66 dementia research group neuropsychological test battery from Singapore's older adult population. Asian Journal of Psychiatry. 2020;51:102019.

2 Ouvrard C, Berr C, Meillon C, Ribet C, Goldberg M, Zins M, et al. Norms for standard neuropsychological tests from the French CONSTANCES cohort. European Journal of Neurology. 2019;26(5):786–93.

3 Luck T, Pabst A, Rodriguez FS, Schroeter ML, Witte V, Hinz A, et al. Age-, sex-, and education-specific norms for an extended CERAD neuropsychological Assessment Battery-results from the population-based LIFE-Adult-Study. Neuropsychology. 2018;32(4):461–75.

4 Sakuma N, Ura C, Miyamae F, Inagaki H, Ito K, Niikawa H, et al. Distribution of Mini-Mental State Examination scores among urban community-dwelling older adults in Japan. International journal of geriatric psychiatry. 2017;32(7):718–25.

5 Alenius M, Koskinen S, Hallikainen I, Ngandu T, Lipsanen J, Sainio P, et al. Cognitive Performance among Cognitively Healthy Adults Aged 30–100 Years. Dementia and Geriatric Cognitive Disorders Extra. 2019;9(1):11–23.

6 Kim H, Au R, Thomas RJ, Yun C-H, Lee SK, Han C, et al. Cognitive performance norms from the Korean genome and epidemiology study (KoGES). International Psychogeriatrics. 2017;29(11):1909–24.

7 Lipnicki DM, Crawford JD, Dutta R, Thalamuthu A, Kochan NA, Andrews G, et al. Age-related cognitive decline and associations with sex, education and apolipoprotein E genotype across ethnocultural groups and geographic regions: a collaborative cohort study. PLoS Medicine. 2017;14(3). DOI: 10.1371/journal.pmed.1002261

8 Prince M, Wimo A, Guerchet M, Ali GC, Wu Y, Prina AM. World Alzheimer Report 2015: The global impact of dementia. An analysis of prevalence, incidence, costs and trends. London: Alzheimer's Disease International; 2015.

9 Harada CN, Natelson Love C. M, Triebel K. Normal Cognitive Aging. Clinics in geriatric medicine. 2013;29(4):737–52.

10 Formanek T, Kagstrom A, Winkler P, Cermakova P. Differences in cognitive performance and cognitive decline across European regions: a population-based prospective cohort study. European Psychiatry. 2019;58:80–6.

11 Chary E, Amieva H, Pérès K, Orgogozo J-M, Dartigues J-F, Jacqmin-Gadda H. Short-versus long-term prediction of dementia among subjects with low and high educational levels. Alzheimer's and Dementia. 2013;9(5):562–71.

12 Abbott RA, Skirrow C, Jokisch M, Timmers M, Streffer J, van Nueten L, et al. Normative data from linear and nonlinear quantile regression in CANTAB: Cognition in mid-to-late life in an

epidemiological sample. Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring. 2019;11:36–44.

13 Meng X, D'Arcy C. Education and dementia in the context of the cognitive reserve hypothesis: A systematic review with meta-analyses and qualitative analyses. PLoS ONE. 2012;7(6). DOI: 10.1371/journal.pone.0038268

14 Stern Y, Arenaza-Urquijo E, Bartrés-Faz D, Belleville S, Cantilon M, Chetelat G, et al. Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. Alzheimer's & Dementia. 2020;16(9):1305–11.

15 Weber D, Skirbekk V, Freund I, Herlitz A. The changing face of cognitive gender differences in Europe. Proceedings of the National Academy of Sciences of the United States of America. 2014 Aug;111(32):11673–8.

16 Elosúa MR, Ciudad MJ, Contreras MJ. Gender differences in verbal and visuospatial working memory tasks in patients with mild cognitive impairment and Alzheimer disease. Dementia and Geriatric Cognitive Disorders Extra. 2017;7(1):101–8.

17 Tensil M, Hessler JB, Gutsmiedl M, Riedl L, Grimmer T, Diehl-Schmid J. Sex Differences in Neuropsychological Test Performance in Alzheimer's Disease and the Influence of the ApoE Genotype. Alzheimer Disease and Associated Disorders. 2018;32(2):145–9.

18 Finland OS of. Population structure [e-publication]. ISSN=1797-5395. Annual Review 2018. [Internet]. 2019;2020(Feb 4,). Available from: http://www.stat.fi/til/vaerak/2018/02/vaerak_2018_02_2019-06-19_tie_001_en.html

19 Finland OS of. Population projection [e-publication]. The decline in the birth rate is reflected in the population development of areas [Internet]. 2019;2019(Oct 25,). Available from: http://www.stat.fi/til/vaenn/2019/vaenn_2019_2019-09-30_tie_001_en.html

20 Finland OS of. Educational structure of population [e-publication]. ISSN=2242-2919. [Internet]. 2020;2020(Nov 8,). Available from: http://www.stat.fi/til/vkour/tau_en.html

21 Hänninen T, Pulliainen V, Salo J, Hokkanen L, Erkinjuntti T, Koivisto K, et al. [Cognitive tests in diagnosing memory disorders and early dementia: CERAD-neuropsychological battery (in Finnish)]. Suomen Lääkärilehti. 1999;54:1967–75.

22 Antikainen A, Kauppila J. Educational generations and the futures of adult education: a Nordic experience. null. 2002;21(3):209–19.

23 Leijola L. The Education System in Finland - Development and Equality. 2004

24 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. 1 Neurology. 1989 Sep;39(9):1159–65.

25 Chandler MJ, Lacritz LH, Hynan LS, Barnard HD, Allen G, Deschner M, et al. A total score for the CERAD neuropsychological battery. Neurology. 2005;65(1):102–6.

26 Seo EH, Lee DY, Lee JH, Choo IH, Kim JW, Kim SG, et al. Total scores of the CERAD neuropsychological assessment battery: Validation for mild cognitive impairment and dementia patients with diverse etiologies. American Journal of Geriatric Psychiatry. 2010;18(9):801–9.

27 Han JY, Seo EH, Yi D, Sohn BK, Choe YM, Byun MS, et al. A normative study of total scores of the CERAD neuropsychological assessment battery in an educationally diverse elderly population. International Psychogeriatrics. 2014;26(11):1897–904.

28 Wolfsgruber S, Jessen F, Wiese B, Stein J, Bickel H, Mösch E, et al. The CERAD neuropsychological assessment battery total score detects and predicts alzheimer disease dementia with high diagnostic accuracy. American Journal of Geriatric Psychiatry. 2014;22(10):1017–28.

29 Rossetti HC, Munro Cullum C, Hynan LS, Lacritz LH. The CERAD Neuropsychologic Battery Total Score and the Progression of Alzheimer Disease. Alzheimer Disease & Associated Disorders. 2010;24(2). Available from: https://journals.lww.com/alzheimerjournal/Fulltext/2010/04000/The_CERAD_Neuropsychologic_Battery_Total_Score_and.5.aspx

30 Paajanen T, Hänninen T, Tunnard C, Hallikainen M, Mecocci P, Sobow T, et al. CERAD Neuropsychological Compound Scores are Accurate in Detecting Prodromal Alzheimer's Disease: A Prospective AddNeuroMed Study. Journal of Alzheimer's Disease. 2014 Feb;39(3):679–90.

31 Suhonen J, Alhainen K, Eloniemi-Sulkava U, Juhela P, Juva K, Löppönen M, et al. Hyvät hoitokäytännöt etenevien muistisairauksien kaikissa vaiheissa. Lääkärilehti. 2008;63:9–22.

32 Pulliainen T, Hänninen T, Hokkanen L, al. et. [Norms for use of the CERAD test battery in Finland]. Suomen Lääkärilehti. 2007;12:1235–41.

33 Hänninen T, Pulliainen V, Sotaniemi M, Hokkanen L, Salo J, Hietanen M, et al. [Early detection of cognitive changes in memory diseases: new cut-off scores for the Finnish version of CERAD neuropsychological battery]. Duodecim; laaketieteellinen aikakauskirja. 2010;126(17):2013–21.

34 Sotaniemi M, Pulliainen V, Hokkanen L, Pirttilä T, Hallikainen I, Soininen H, et al. CERAD-neuropsychological battery in screening mild Alzheimer's disease. Acta Neurologica Scandinavica. 2012;125(1):16–23.

35 Commission E. The 2021 Ageing Report Economic and Budgetary Projections for the EU Member States 2019-2070. Part III Statistical Annex – CROSS-COUNTRY TABLES, Table III.1.14: Very elderly population (80 and over) as % of elderly population [Internet]. 2021;2021(Aug 24,). Available from: https://ec.europa.eu/info/publications/2021-ageing-report-economic-and-budgetary-projections-eu-member-states-2019-2070_en

36 Ngandu T, Lehtisalo J, Levälahti E, Laatikainen T, Lindström J, Peltonen M, et al. Recruitment and baseline characteristics of participants in the finnish geriatric intervention study to prevent cognitive impairment and disability (FINGER)—A randomized controlled lifestyle trial. International Journal of Environmental Research and Public Health. 2014;11(9):9345–60.

37 Duodecim W group set up by the FMS. [Diagnosis, symptom evaluation and disease follow-up of memory disease. Current Care Guidelines.][Muistisairauden diagnostiikka, oireiden arviointi ja sairauden seuranta] [Internet]. 2016;2019(Apr 7). Available from: http://www.kaypahoito.fi/web/kh/suositukset/suositus

38 McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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. 1984;34(7):939–44.

39 Association AP. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV. 4th edition. Washington: American Psychiatric Association; 1994.

40 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(3):189–98.

41 Auer S, Reisberg B. The GDS/FAST staging system. International Psychogeriatrics. 1997;9(SUPPL. 1):167–71.

42 Morris JC. Clinical Dementia Rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. International Psychogeriatrics. 1997;9(SUPPL. 1):173–6.

43 Karrasch M, Laatu S, Ellfolk U, Marttila R, Martikainen K. Education-corrected CERAD identifies MCI and dementia in Parkinson's disease. Acta Neurologica Scandinavica. 2015;131(4):219–24.

Figure Legends

Fig. 1. Sampling of patients with a diagnosis of Alzheimer's disease (AD). AD, Alzheimer's disease; CERAD-nb, Consortium to Establish a Register for Alzheimer's Disease neuropsychological battery; LHR, local health record; NHR, national health register; PSP, private sector provider.

Fig 1

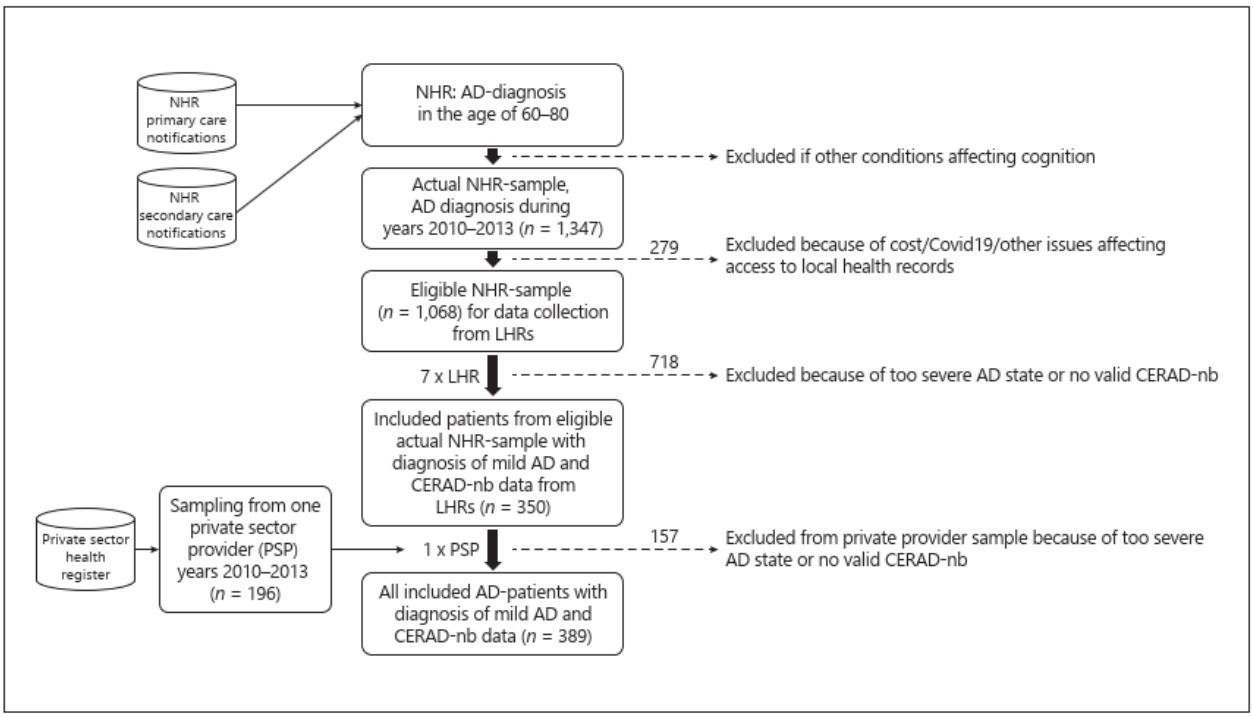
NHR primary care notifications
NHR secondary care notifications
NHR: AD-diagnosis in the age of 60–80
Excluded if other conditions affecting cognition
Actual NHR-sample, AD diagnosis during years 2010–2013 (n = 1,347)
279
Excluded because of cost/Covid19/other issues affecting access to local health records
Eligible NHR-sample (n = 1,068) for data collection from LHRs
7 x LHR
718
Excluded because of too severe AD state or no valid CERAD-nb
Included patients from eligible actual NHR-sample with diagnosis of mild AD and CERAD-nb data from LHRs (n = 350)
Private sector health register
Sampling from one private sector provider (PSP) years 2010–2013 (n = 196)
1 x PSP
157
Excluded from private provider sample because of too severe AD state or no valid CERAD-nb
All included AD-patients with diagnosis of mild AD and CERAD-nb data (n = 389)

Table 1. Study population.

		Total	Controls	AD group
	Total, n (%)	2369 (100)	1980 (100)	389 (100)
Gender [a)]	Women, n (%)	1247 (53)	1022 (52)	225 (58)
	Men, n (%)	1122 (47)	958 (48)	164 (42)
Education [b)]	Missing education information n (%)	20 (1)		20 (5)
	Lowest level, n (%)	389 (16)	222 (11)	167 (43)
	Middle level, n (%)	1152 (49)	1027 (52)	125 (32)
	Highest level, n (%)	808 (34)	731 (37)	77 (20)
Age group [c)]	60–69 years, n (%)	1158 (49)	1093 (55)	65 (17)
	70–74 years, , n (%)	675 (28)	555 (28)	120 (31)
	75–81 years, n (%)	536 (23)	332 (17)	204 (52)
Cohort [d)]	born 1930–1934, n (%)	364 (15)	224 (11)	140 (36)
	born 1935–1939, n (%)	683 (29)	535 (27)	148 (38)
	born 1940–1944, n (%)	698 (29)	630 (32)	68 (17)
	born 1945–1949, n (%)	618 (26)	591 (30)	27 (7)
	born 1950–1953, n (%)	6 (0)		6 (2)
Generation [e)]	War generation	461 (19)	288 (15)	173 (44)
	Stuctural change generation	1908 (81)	1692 (85)	216 (56)

Abbreviations: controls, the non-demented comparison group; AD, Alzheimer's disease; Lowest level, <= 7 years (former primary school; Middle level, ~8–10 years (e.g. middle/ technical/ trade/ vocational school) or matriculation examination without any other degree; Highest level, >10 years with degree from college or university. War generation, born 1935 or before, with scant education; Structural change generation, born 1936–1955, with growing educational opportunities.

[a)] Gender: no difference (p =.153) between controls and AD group.

[b)] Education: controls more educated than AD group (p <.001).

[c)] Age groups: controls younger than AD group (p <.001).

[d)] Cohorts of controls younger than AD group (p <.001).

[e)] Generations differed between controls and AD group (p <.001).

Table 2. Statistical parameters CERAD-nb in 1980 controls and 389 AD-patients. Group difference was tested with univariate GLM controlling for age, education, and gender.

	TOTAL			
	Controls Mean (SD) (n=1980)	AD group Mean (SD) (n=389)	Significance of Group Difference	Partial Eta Squared η_p^2 for Group Difference ($p<.05$)
Verbal Fluency, animal category	23.1 (6.0)	13.9 (4.8)	p<.001	.157
15-item Boston Naming Test	13.3 (1.7)	10.6 (2.9)	p<.001	.077
MMSE	27.6 (1.9)	23.8 (2.4)	p<.001	.207
Word List Memory	20.7 (3.6)	13.0 (3.6)	p<.001	.269
Constructional Praxis Copy	9.9 (1.4)	9.2 (1.9)		
Word List Recall	6.8 (1.9)	2.4 (1.9)	p<.001	.305
Word List Savings (%)	83.7 (16.7)	44.7 (30.7)	p<.001	.239
Word List Recognition	19.1 (1.2)	16.0 (2.6)	p<.001	.238
Word List Recognition %	95.6 (6.1)	80.2 (12.8)	p<.001	.238
Verbal memory total recall	27.1 (2.2)	21.3 (3.3)	p<.001	.307
Constructional Praxis Recall	8.9 (2.1)	4.7 (3.3)	p<.001	.169
Constructional Praxis Savings (%)	88.6 (16.1)	49.5 (32.9)	p<.001	.239
Clock Drawing	5.1 (1.3)	4.2 (1.6)	p<.001	.019
Chandler Total	80.3 (9.4)	55.7 (11.0)	p<.001	.314
Seo Total	89.3 (10.4)	60.6 (12.7)	p<.001	.332

Abbreviations: AD, Alzheimer's disease; controls, the non-demented comparison group

Table 3a. Proportion of variance in cognitive performance explained by demographic factors in 1980 controls. Effect sizes as Eta Squared η^2 ($p < .05$) for each separate model of Education, Age and Gender; and as Partial Eta Squared η_p^2 for combined model of Education, Age, and Gender.

	Eta Squared η^2 ($p<.05$)	Eta Squared η^2 ($p<.05$)	Eta Squared η^2 ($p<.05$)	**Partial Eta Squared η_p^2** ($p<.05$) for combination of factors Education, Age, Gender							
	Model Education	Model Age	Model Gender	Model Education, Age, Gender	Interaction Education* Age*Gender	Interaction Education* Age	Interaction Education* Gender	Interaction Age*Gender	Education	Age	Gender
Verbal Fluency, animal category	.083 ***	.017 ***		.096 ***					.057 ***	.011 ***	
15-item Boston Naming Test	.095 ***	.029 ***	.055 ***	.166 ***			.004 *		.074 ***	.016 ***	.033 ***
MMSE	.077 ***	.023 ***		.102 ***		.008 **			.059 ***	.017 ***	
Word List Memory	.062 ***	.031 ***	.037 ***	.133 ***					.051 ***	.017 ***	.032 ***
Constructional Praxis Copy	.071 ***	.031 ***		.098 ***					.047 ***	.018 ***	
Word List Recall	.044 ***	.025 ***	.012 ***	.081 ***					.031 ***	.015 ***	.011 ***
Word List Savings (%)	.014 ***	.007 **		.027 ***					.008 **		
Word List Recognition	.020 ***	.019 ***	.012 ***	.057 ***					.014 ***	.013 ***	.015 ***
Word List Recognition (%)	.020 ***	.019 ***	.012 ***	.057 ***					.014 ***	.013 ***	.015 ***
Verbal Memory total recall	.039 ***	.033 ***	.023 ***	.100 ***					.031 ***	.024 ***	.024 ***
Constructional Praxis Recall	.107 ***	.032 ***	.006 ***	.131 ***					.075 ***	.016 ***	
Constructional Praxis Savings (%)	.049 ***	.012 ***	.006 ***	.065 ***					.033 ***	.007 **	
Clock Drawing	.006 **	.017 ***	.004 **	.030 ***					.004 *	.010 ***	.004 **
Chandler Total	.129 ***	.050 ***	.004 **	.171 ***					.097 ***	.031 ***	.009 ***
Seo Total	.150 ***	.056 ***		.193 ***					.114 ***	.035 ***	.006 ***

Abbreviations: controls, the non-demented comparison group. Symbols: * $p<.05$; ** $p<.01$; *** $p<.001$

Table 3b. Proportion of variance in cognitive performance explained by demographic factors in 389 patients with AD diagnosis. Effect sizes as Eta Squared η^2 ($p < .05$) for each separate model of Education, Age and Gender; and as Partial Eta Squared η_p^2 for combined model of Education, Age, and Gender.

	Eta Squared η^2 ($p<.05$)	Eta Squared η^2 ($p<.05$)	Eta Squared η^2 ($p<.05$)	**Partial Eta Squared η_p^2** ($p<.05$) for combination of factors Education, Age, Gender							
	Model Education	Model Age	Model Gender	Model Education, Age, Gender	Interaction Education* Age*Gender	Interaction Education* Age	Interaction Education* Gender	Interaction Age*Gender	Education	Age	Gender
Verbal Fluency, animal category	.030 **			.093 **				.017 **			
15-item Boston Naming Test	.076 ***	.081 ***	.094 ***	.246 ***					.034 **	.041 ***	.040 ***
MMSE	.098 ***		.014*	.140 ***				.018 *	.054 ***		.017 *
Word List Memory	.083 ***	.019 *		.148 ***				.020 *	.043 ***	.022 *	
Constructional Praxis Copy	.097 ***	.026 **		.178 ***					.051 ***		
Word List Recall	.043 **		.013 *	.120 **		.039 *	.028 *	.023 *	.046 **	.025 *	.031 *
Word List Savings (%)	.021 *		.011 *								
Word List Recognition	.048 ***			.102 *	.032 *			.019 *	.038 **		.011 *
Word List Recognition (%)	.048 ***			.102 *	.032 *			.019 *	.038 **		.011 *
Verbal Memory total recall	.068 ***			.129 **	.033 *			.030 **	.056 ***		.014 *
Constructional Praxis Recall	.097 ***	.032 **	.019 *	.161 ***					.054 ***	.023 *	.012 *
Constructional Praxis Savings (%)	.054 ***	.022 *	.012 *	.110 **					.032 **	.026 **	.017 *
Clock Drawing	.028 *	.027 **									
Chandler Total	.147 ***	.051 ***	.018 *	.260 ***		.050 *			.071 ***	.024 *	.018 *
Seo Total	.167 ***	.054 ***	.023 **	.279 ***		.050 *			.085 ***	.026 *	.024 *

Abbreviations: AD, Alzheimer's disease. Symbols: * $p<.05$; **$p<.01$; *** $p<.001$

Table 4a. Cognitive performance by education (means and (standard deviations)) and differences for CERAD-nb in 1980 controls and in 389 AD-patients. Pairwise group differences were tested post-hoc following a significant univariate GLM controlling for age and gender.

	Controls				**AD-group**			
	Lowest level (n=222)	Middle level (n=1027)	Highest level (n=731)	Differences between education groups	Lowest level (n=167)	Middle level (n=125)	Highest level (n=77)	Differences between education groups
Verbal Fluency, animal category	20.7 (5.5)	22.0 (5.6)	25.3 (6.1)	# (p<.01), ## (p<.001), + (p<.001)	13.1 (4.7)	14.5 (4.5)	14.9 (5.2)	## (p<.05)
15-item Boston Naming Test	12.3 (2.1)	13.0 (1.8)	13.9 (1.2)	# (p<.001), ## (p<.001), + (p<.001)	9.8 (2.9)	11.0 (2.6)	11.8 (2.8)	# (p<.01), ## (p<.001)
MMSE	26.7 (2.2)	27.4 (1.9)	28.2 (1.5)	# (p<.001), ## (p<.001), + (p<.001)	23.1 (2.2)	24.0 (2.4)	25.1 (2.3)	# (p<.01), ## (p<.001), + (p<.01)
Word List Memory	19.6 (3.6)	20.1 (3.5)	21.8 (3.4)	## (p<.001), + (p<.001)	11.9 (3.2)	13.6 (3.5)	14.5 (3.9)	# (p<.001), ## (p<.001)
Constructional Praxis Copy	9.3 (1.6)	9.7 (1.4)	10.3 (1.1)	# (p<.001), ## (p<.001), + (p<.001)	8.7 (1.8)	9.5 (1.7)	10.1 (1.4)	# (p<.01), ## (p<.001)
Word List Recall	6.3 (1.8)	6.5 (1.9)	7.3 (1.8)	## (p<.001), + (p<.001)	2.1 (1.7)	2.5 (1.7)	3.1 (2.0)	## (p<.001)
Word List Savings (%)	81.5 (16.4)	82.4 (17.6)	86.3 (15.2)	## (p<.001), + (p<.001)	40.4 (32.3)	45.3 (29.7)	52.0 (28.2)	## (p<.01)
Word List Recognition	18.9 (1.4)	19.0 (1.3)	19.3 (1.0)	## (p<.001), + (p<.001)	15.5 (2.5)	16.0 (2.6)	17.0 (2.5)	## (p<.001), + (p<.05)
Word List Recognition %	94.6 (6.8)	95.0 (6.4)	96.7 (5.2)	## (p<.001), + (p<.001)	77.7 (12.5)	80.2 (12.9)	85.1 (12.3)	## (p<.001), + (p<.05)
Verbal memory total recall	26.5 (2.3)	26.8 (2.3)	27.6 (2.0)	## (p<.001), + (p<.001)	20.4 (3.0)	21.4 (3.3)	22.7 (3.5)	## (p<.001), + (p<.05)
Constructional Praxis Recall	8.0 (2.1)	8.5 (2.1)	9.8 (1.7)	# (p<.01), ## (p<.001), + (p<.001)	3.9 (3.0)	4.9 (3.3)	6.4 (3.2)	## (p<.001), + (p<.01)
Constructional Praxis Savings (%)	85.2 (17.2)	86.0 (17.3)	93.2 (12.6)	## (p<.001), + (p<.001)	44.3 (31.8)	50.3 (33.1)	62.9 (29.8)	## (p<.001), + (p<.05)
Clock Drawing	4.9 (1.5)	5.1 (1.3)	5.2 (1.2)	# (p<.05), ## (p<.01)	4.0 (1.7)	4.5 (1.6)	4.4 (1.4)	## (p<.05)
Chandler Total	75.5 (9.9)	78.4 (9.0)	84.6 (7.9)	# (p<.001), ## (p<.001), + (p<.001)	51.2 (10.4)	57.9 (9.2)	61.7 (11.1)	# (p<.01), ## (p<.001)
Seo Total	83.6 (11.0)	86.8 (9.9)	94.4 (8.6)	# (p<.001), ## (p<.001), + (p<.001)	55.2 (11.8)	63.0 (10.7)	68.3 (12.6)	# (p<.001), ## (p<.001), + (p<.05)

Abbreviations: controls, the non-demented comparison group; Lowest level, <= 7 years (former primary school; Middle level, ~8–10 years (e.g., middle/ technical/ trade/ vocational school) or matriculation examination without any other degree; Highest level, >10 years with degree from college or university.
Symbols: # = Lowest level< Middle level, ## = Lowest level< Highest level, + = Middle level < Highest Level, " = Highest < Lowest, "" = Highest < Middle

Table 4b. Cognitive performance by age (means and (standard deviations)) and differences for CERAD-nb in 1980 controls and in 389 AD-patients. Pairwise group differences were tested post-hoc following a significant univariate GLM controlling for education and gender.

	Controls				AD-group			
	60–69 years (n=1093)	70–74 years (n=555)	75–81 years (n=332)	Differences between age groups	60–69 years (n=65)	70–74 years (n=120)	75–81 years (n=204)	Differences between age groups
Verbal Fluency, animal category	23.7 (6.1)	22.8 (6.0)	21.5 (5.3)	# (p<.01), ## (p<.001), + (p<.05)	14.8 (5.0)	14.2 (4.7)	13.5 (4.7)	
15-item Boston Naming Test	13.5 (1.6)	13.1 (1.8)	12.8 (2.0)	# (p<.05), ## (p<.001), + (p<.001)	11.9 (2.3)	11.2 (2.6)	9.9 (3.0)	# (p<.001), ##(p<.001)
MMSE	27.8 (1.7)	27.5 (1.9)	27.1 (2.1)	# (p<.01), ## (p<.001), + (p<.001)	24.2 (2.4)	23.8 (2.5)	23.7 (2.4)	
Word List Memory	21.2 (3.4)	20.2 (3.6)	19.7 (3.8)	## (p<.001), + (p<.001)	14.1 (4.2)	13.1 (3.3)	12.7 (3.6)	## (p<.05)
Constructional Praxis Copy	10.1 (1.3)	9.8 (1.4)	9.5 (1.5)	# (p<.01), ## (p<.001), + (p<.001)	9.9 (1.5)	9.1 (1.9)	9.1 (1.9)	## (p<.05), + (p<.05)
Word List Recall	7.0 (1.8)	6.7 (1.9)	6.2 (1.9)	# (p<.01), ## (p<.001), + (p<.01)	2.7 (1.9)	2.4 (1.9)	2.4 (1.8)	
Word List Savings (%)	84.8 (15.8)	83.1 (17.9)	81.2 (17.3)	## (p<.001)	47.8 (29.9)	42.4 (31.6)	45.1 (30.5)	
Word List Recognition	19.2 (1.1)	19.1 (1.2)	18.8 (1.5)	# (p<.001), ## (p<.001)	16.3 (2.7)	15.9 (2.5)	16.1 (2.5)	
Word List Recognition %	96.1 (5.6)	95.6 (6.1)	93.8 (7.3)	# (p<.001), ## (p<.001)	81.6 (13.6)	79.4 (12.7)	80.3 (12.6)	
Verbal memory total recall	27.3 (2.1)	27.0 (2.2)	26.2 (2.5)	# (p<.001), ## (p<.001), + (p<.01)	21.9 (3.7)	21.1 (3.4)	21.2 (3.1)	
Constructional Praxis Recall	9.2 (1.9)	8.7 (2.2)	8.3 (2.2)	# (p<.05), ## (p<.001), + (p<.01)	5.6 (3.5)	5.1 (3.2)	4.1 (3.1)	# (p<.05), ## (p<.01)
Constructional Praxis Savings (%)	90.1 (14.8)	87.2 (17.2)	85.7 (17.6)	## (p<.001), + (p<.01)	55.1 (33.6)	54.4 (32.6)	45.0 (32.4)	# (p<.05)
Clock Drawing	5.3 (1.1)	5.0 (1.3)	4.9 (1.4)	## (p<.001), + (p<.001)	4.5 (1.7)	4.5 (1.5)	4.0 (1.6)	# (p<.05)
Chandler Total	82.1 (8.7)	79.2 (9.4)	76.6 (10.0)	# (p<.001), ## (p<.001), + (p<.001)	60.7 (10.1)	55.9 (10.3)	53.8 (11.2)	# (p<.05), ## (p<.01)
Seo Total	91.3 (9.6)	87.9 (10.5)	84.8 (11.1)	# (p<.001), ## (p<.001), + (p<.001)	66.5 (11.8)	61.0 (11.9)	58.3 (13.0)	# (p<.05), ## (p<.001)

Abbreviations: AD, Alzheimer's disease; controls, the non-demented comparison group
Symbols: # = 75-81 years < 70-74 years, ## = 75-81 years < 60-69 years, + =70-74 years < 60-69 years
“ = 60-69 years < 75-81 years, “” = 60-69 years < 70-74 years

Table 4c. Cognitive performance by gender (means and (standard deviations)) and differences for CERAD-nb in 1980 controls and in 389 AD-patients. Pairwise group differences were tested following a significant univariate GLM controlling for education and age.

	Controls			AD-group		
	Women (n=1022)	Men (n=958)	Difference between genders	Women (n=225)	Men (n=164)	Difference between genders
Verbal Fluency, animal category	22.8 (5.8)	23.3 (6.2)		13.6 (4.5)	14.4 (5.1)	
15-item Boston Naming Test	12.9 (1.8)	13.7 (1.5)	m (p<.001)	9.8 (2.9)	11.6 (2.5)	m (p<.001)
MMSE	27.7 (1.9)	27.6 (1.8)		23.6 (2.3)	24.1 (2.5)	m (p<.05)
Word List Memory	21.3 (3.4)	20.0 (3.6)	w (p<.001)	13.1 (3.6)	12.9 (3.7)	
Constructional Praxis Copy	9.9 (1.3)	9.9 (1.4)		9.1 (2.0)	9.4 (1.7)	
Word List Recall	7.0 (1.8)	6.6 (1.9)	w (p<.001)	2.2 (1.7)	2.7 (2.0)	m (p<.05)
Word List Savings (%)	84.4 (16.3)	83.1 (17.1)		42.0 (29.4)	48.5 (32.0)	m (p<.05)
Word List Recognition	19.2 (1.1)	19.0 (1.3)	w (p<.001)	15.8 (2.5)	16.3 (2.6)	
Word List Recognition %	96.2 (5.6)	94.9 (6.5)	w (p<.001)	79.2 (12.5)	81.7 (13.1)	
Verbal memory total recall	27.4 (2.0)	26.7 (2.3)	w (p<.001)	21.0 (3.2)	21.6 (3.4)	
Constructional Praxis Recall	8.7 (2.1)	9.1 (2.1)	m (p<.001)	4.3 (3.2)	5.1 (3.3)	m (p<.05)
Constructional Praxis Savings (%)	87.3 (16.7)	89.9 (15.3)	m (p<.001)	46.4 (32.6)	53.6 (33.0)	m (p<.05)
Clock Drawing	5.2 (1.2)	5.1 (1.3)	w (p<.01)	4.2 (1.6)	4.4 (1.6)	
Chandler Total	80.9 (9.2)	79.7 (9.5)	w (p<.01)	54.3 (11.0)	57.3 (10.8)	m (p<.05)
Seo Total	89.7 (10.2)	88.8 (10.6)		58.8 (12.7)	62.6 (12.5)	m (p<.01)

Abbreviations: AD, Alzheimer's disease; controls, the non-demented comparison group
Symbols: w = women better than men, m = men better than women

Table 5. Cut-offs for CERAD-nb including 2010-cut-offs [32] and new potential values according to the best Youden's J using uncorrected and corrected cut-off values for 389 patients with AD diagnosis and 1980 controls, and for education, age, and gender groups. Dark grey indicates 2010 cut-offs and related false negative and positive calculations with those cut-offs in the present sample. Medium grey indicates new uncorrected cut-offs and related false negative and positive calculations. Light grey indicates false negative and false positive-rates for using uncorrected cut off-values (instead of education-corrected cut-offs) for participants with the highest level of education.

Cerad measure	**Uncorrected cut-off values**								**Corrected cut-off values**											
									Education			At the highest education level				Age			Gender	
	2010 FalseN	2010 FalseP	2010 cut-offs	Cut-off	**AUC**	Youd	FalseN	FalseP	Lowest level	Middle level	Highest level	FalseN	FalseP	Not used FalseN	Not used FalseP	60–69 years	70–74 years	75–81 years	W	M
Verbal fluency, animal category	0.346	0.092	< 16	19	**0.892**	0.627	0.157	0.217	17	18	20	0.107	0.164	0.187	0.123	19	19	18	19	19
15-item Boston Naming Test	0.441	0.154	< 12	13	**0.778**	0.427	0.303	0.270	11	13	13	0.480	0.124	0.480	0.124	13	13	11	12	13
MMSE [a)]	(0.516)	(0.028)	(24)	26	**0.882**	0.607	0.259	0.134	26	26	27	0.307	0.116	0.480	0.056	26	27	26	26	26
Word List Memory	0.162	0.131	< 17	17	**0.926**	0.707	0.162	0.131	16	18	19	0.160	0.140	0.267	0.070	18	17	16	18	17
Word List Recall	0.138	0.113	< 5	5	**0.937**	0.749	0.138	0.113	5	5	6	0.120	0.166	0.227	0.074	5	5	5	5	5
Word List Savings (%)	0.165	0.282	< 75	68	**0.855**	0.609	0.227	0.164	61	61	73	0.227	0.163	0.280	0.123	68	68	61	68	68
Word List Recognition			-	19	**0.863**	0.584	0.189	0.227	19	19	19	0.347	0.171	0.347	0.171	19	19	18	19	19
Word List Recognition %	0.264	0.167	< 90	93	**0.863**	0.584	0.189	0.227	93	93	93	0.347	0.171	0.347	0.171	93	93	88	93	93
Constructional Praxis Copy			-	9	**0.587**	0.185	0.647	0.168	8	8	9	0.855	0.074	0.855	0.074	8	9	7	10	9
Constructional Praxis Recall			-	7	**0.846**	0.545	0.324	0.132	6	7	9	0.280	0.204	0.347	0.164	7	7	6	6	8
Constructional Praxis Savings (%)	0.403	0.067	< 60	74	**0.836**	0.548	0.255	0.197	65	74	84	0.267	0.183	0.387	0.105	87	74	59	74	74
Clock	0.465	0.271	< 5	6	**0.665**	0.276	0.322	0.402	5	6	6	0.293	0.378	0.293	0.378	5	5	6	6	5
Verbal memory total recall	0.362	0.045	< 23	25	**0.926**	0.716	0.159	0.125	25	26	27	0.147	0.234	0.333	0.075	26	25	25	25	25
Chandler Total			-	71	**0.951**	0.755	0.094	0.151	64	68	76	0.049	0.129	0.213	0.052	74	67	65	67	71

Cerad measure	Uncorrected cut-off values								Corrected cut-off values												
									Education			At the highest education level				Age			Gender	
	2010 FalseN	2010 FalseP	2010 cut-offs	Cut-off	**AUC**	Youd	FalseN	FalseP	Lowest level	Middle level	Highest level	FalseN	FalseP	Not used FalseN	Not used FalseP	60–69 years	70–74 years	75–81 years	W	M
Seo Total			-	79	**0.954**	0.757	0.087	0.156	70	78	82	0.115	0.083	0.213	0.055	80	72	69	75	80

Abbreviations: 2010-cut-offs, the Finnish cut-off scores for CERAD-nb published in 2010; 2010 FalseN, False Negative if 2010 cut-off value would have been used in this sample; 2010 FalseP, False Positive if 2010 cut-off value would have been used in this sample; FalseN, False Negative (=1- sensitivity); FalseP, False Positive (=1- specificity). Not used, if cut-off values would not be corrected but would be the same for all. Lowest level, <= 7 years (former primary school); Middle level, ~8–10 years (e.g., middle/ technical/ trade/ vocational school) or matriculation examination without other degree; Highest level, >10 years with degree from college or university. W, Women; M, Men. Youd, Youden's J indicating best combination of sensitivity and specificity.

[a)] MMSE is not part of the 2010 cut-offs, but another recommendation [21] has been given for Finnish cut-off value 24.