Title

Neuropsychological Test Performance and MRI Markers of Dementia Risk

Journal

Alzheimer Disease & Associated Disorders, 33(3)

ISSN

0893-0341

Authors

Walter, Stefan
Dufouil, Carole
Gross, Alden L
et al.

Publication Date

2019-07-01

DOI

10.1097/wad.0000000000000321

Peer reviewed

Published in final edited form as:

Alzheimer Dis Assoc Disord. 2019 ; 33(3): 179–185. doi:10.1097/WAD.0000000000000321.

Neuropsychological Test Performance and MRI Markers of Dementia Risk: Reducing Education Bias

Stefan Walter, PhD[1,2], **Carole Dufouil, PhD**[3], **Alden L. Gross, PhD**[4], **Richard N. Jones, ScD**[5], **Dan Mungas, PhD**[6], **Teresa J. Filshtein, PhD**[2], **Jennifer J. Manly, PhD**[7], **Thalida E. Arpawong, PhD**[8], **M. Maria Glymour, ScD**[2]

[1]Fundación de Investigación Biomedica Hospital Universitario de Getafe, KM 12.5 Crta de Toledo, 28905 Getafe, Madrid, Spain

[2]Department of Epidemiology and Biostatistics, 550 16th Street, University of California, San Francisco 94158

[3]Inserm, Bordeaux Population Health Research Center, UMR 1219, Univ. Bordeaux, ISPED, CIC 1401-EC, F-33000 Bordeaux, France

[4]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[5]Department of Psychiatry & Human Behavior, Department of Neurology, Warren Alpert Medical School, Brown University, Providence, Rhode Island, USA 02906

[6]Davis School of Medicine, University of California, Sacramento 95817, USA

[7]Department of Neurology and the Taub Institute for Research on Alzheimer's Disease and the Aging Brain

[8]Davis School of Gerontology, University of Southern California, 3715 McClintock Avenue, Los Angeles, CA 90089

Abstract

Background: To use neuropsychological assessments for studying the underlying disease processes contributing to dementia, it is crucial that they correspond to MRI-based measures of dementia regardless of educational level.

Methods: French 3 City Dijon MRI study cohort members (n=1,782) with assessments of white matter lesion volume (WMLV), hippocampal volume (HCV), and cerebrospinal fluid volume (CSFV) and 6 waves of neuropsychological assessments over 11 years, including Mini-Mental State Examination (MMSE), plus 5 other tests combined using a Z-score or item-response theory (IRT-cognition). We evaluated, testing interactions, whether education modified associations of MRI markers with intercept or rate of change of MMSE, Z-score composite, or IRT-cognition.

Corresponding Author: M. Maria Glymour, Department of Epidemiology and Biostatistics, 550 16th Street, University of California, San Francisco 94158, mglymour@psg.ucsf.edu.

Disclosures:
The authors report no disclosures.

Results: In linear models, education modified the associations of WMLV and CSFV with MMSE and CSFV and Z score composite. In mixed models, education modified the associations of WMLV and CSFV with level of MMSE and the association of HCV with slope of MMSE. Education also modified the association with CSFV and slope of Z-score composite decline. There was no evidence that education modified associations between MRI measures and level or slope of IRT-cognition.

Conclusion: Longitudinal analysis of correctly scaled neuropsychological assessments may provide unbiased proxies for MRI based measures of dementia risk.

Introduction

Neuropsychological tests offer noninvasive and inexpensive methods to detect elevated dementia risk. Exclusive reliance on biomarker-based measures of Alzheimeŕs disease, the most frequent cause of dementia, is likely to reduce the diversity and size of human studies of the etiology of Alzheimeŕs disease and dementia.[1] Analytic approaches to enhance the validity of non-invasive assessments of disease progression, e.g., neuropsychological assessments, are critical.

Because both progressive biological disease processes and cognitive or brain reserve contribute to dementia diagnoses, it is important to evaluate risk factors for both. Brief neuropsychological assessments and screening tools, such as the Mini Mental State Examination (MMSE), may provide valuable tools for research if these neuropsychological measures correspond with underlying progressive pathological processes similarly across populations of different educational backgrounds. Prior research suggests, however, that education modifies the associations between some neuropsychological tests and pathological burden.[2–6] If a 1-point difference in neuropsychological test performance does not equate to the same magnitude of difference in underlying disease progression for a high or low education individuals, the measure will be a biased tool for assessing this underlying pathology (Figure 1). Similarly, such a bias will compromise etiologic research on any putative risk factor that is correlated with education.

We hypothesized that with optimal neuropsychological measurement and modeling, the correspondence between neuroimaging markers of dementia risk and neuropsychological assessments would be equivalent regardless of education. We examined the relationships of MRI-based biomarkers of brain health with level and change in neuropsychological tests characterized using a single measure (MMSE score), a Z-score composite measure of 4 tests, and an item-response-theory based composite measure (IRT-Cognition).

Methods:

Population:

The 3 City (3C) Dijon study is a population-based cohort of 4,931 French noninstitutionalized individuals (Three-City Study, 2003).[7] A total of 2,763 individuals aged < 80 years were invited to undergo brain MRI between June 1999 and September 2000. Consent rate was high (83%, 2,285 individuals) but because of financial restrictions, only 1,923 MRI scans were performed. Valid MRI measures for all exposure variables were

available for 1,782 participants aged between 65 to 80 years. On average these 1,782 participants were 1.5 years younger, scored higher on all cognitive tests, and were better educated than the entire 3C sample. For example, 46% of MRI participants completed up to upper primary school (8 years or less), while 58% of the full sample completed up to upper primary school. The cognitive status of all participants was assessed up to 6 times over 11 years of follow-up. MRIs were completed at baseline.

The 3C study was approved by the Institutional Review Board at Kremlin-Bicêtre University Medical Center, Paris, France. All participants gave their written, informed consent to participation.

Measures:

MRI acquisition details have been previously reported.[8,9] White Matter Lesion Volume (WMLV) was estimated from neuroimaging data with an automatic image processing method and Hippocampal Volume (HCV) and Cerebral Spinal Fluid Volume (CSFV) were estimated with voxel-based morphometry. We normalized CSFV, HCV, and WMLV by residualizing against Intracranial Volume (ICV) and rescaling to the standard normal distribution. To simplify comparisons, we use Z-score transformed versions of the MRI measures.

Education was assessed in 3C as an ordinal variable and dichotomized at the median (upper primary school [8 years] or less versus more than 8 years).

At each wave, cognitive testing included the MMSE, Benton visual recognition test, Trail Making Tests A and B, and verbal fluency (Isaacs Set Test). The MMSE is a screening test that assesses global cognition (working memory, language and praxis, orientation, memory, and attention).[10] Despite the documented limitations of the MMSE, we used this as a reference test because MMSE or variants are extremely common tools for assessing dementia risk in epidemiological research studies. The Benton Visual Retention Test assesses nonverbal memory and is associated with construction and design copying tasks[11]; the Trail Making Tests[12] assess attention, visuomotor tracking and speed, divided attention, and cognitive flexibility[13], and the Isaacs Set Test reflects both literacy and the ability to organize thinking by clustering words from cities, fruits, vegetables, and colors.[14]

Analysis

We evaluated associations between MRI measures and three cognitive measures. To place all 3 cognitive outcomes on the same scale, we Z-score transformed MMSE using its baseline mean and standard deviation. We also created Z-score composite and IRT-based composite measures of cognition. To create the Z-score composite, each measure (including MMSE) was first Z-score transformed using its baseline mean and standard deviation. We then averaged the Z-scores for the 6 instruments and rescaled to a standard normal distribution in the analytic sample.

To estimate the IRT-Cognition measures for each participant, we used confirmatory factor analysis as implemented in the lavaan R package (R Version 3.2.3) to estimate a two-parameter IRT model at each assessment wave.[15,16] The model assumes a continuous

underlying latent variable (general cognition) and performance on any specific cognitive test (item) reflects the influence of the latent variable and item-specific residuals.[17] The two-parameter IRT model can be written for each item as a cumulative normal probability function with an item-specific intercept (to distinguish simple from difficult items) and a slope parameter that describes the influence of the underlying latent variable on the probability of successfully answering that item.[17]

Following the approach described by Gross et al.[18,19], the continuous cognitive assessment scores for each test were subdivided into up to 5 categories of approximately equal sample size and treated as ordinal endogenous variables for the estimation model. The latent IRT-Cognition (general cognition) measure was set to have unit variance at baseline. In addition, we constrained factor loadings and test thresholds to be the same across visits while freeing subsequent means and variances of the latent variables at each study visit after baseline. Change in cognitive performance over the course of follow-up is thus reflected in the levels of the predicted latent variable at each study visit.

Missingness in single variable responses was addressed using multiple imputation (with 10 imputed data sets) under the missing at random assumption, such that diagonally weighted least squares (WLSMV) estimation on complete data could be used in the estimation of the IRT model.[20] The final IRT-Cognition measure was scaled to have a mean of 0 and standard deviation of 1 at baseline in the analytic sample.

Next, we evaluated whether education modified the associations of WMLV, HCV, and CSFV (all residualized against ICV) in linear regression models predicting baseline cognitive scores. These models were adjusted for age at first cognitive assessment and sex, and included education, one MRI measure, and the interaction of education and the MRI measure as predictors. We estimated nine linear regression models: one for each of the three MRI measures, for each cognitive outcome (MMSE, Z-score, and IRT-Cognition). We next used linear age and sex adjusted growth curve models with individual-level random intercepts and slopes to compare whether the associations of each MRI measure with either baseline (intercept) or rate of change (slope) in the cognitive outcomes (MMSE, Z-score, or IRT-Cognition) were modified by education. We centered follow-up time in these models at 3.5 years, to avoid estimating intercepts at the extremes of follow-up. Effect modification was tested with an interaction term between education and each MRI measure.

To express the magnitude of effect modification (potential education bias) in meaningful terms, we used 100 times the ratio of the interaction term to the main effect of the MRI measure. If the MRI and neuropsychological assessment had equivalent correspondence for low and high education individuals, this ratio would be zero; a value of 100 indicates that the effect modification (interaction) in one education group is as large as the main effect in the other education group. To avoid over-interpreting trivially small biases that met statistical significance criteria or ignoring large biases that were imprecisely estimated, we defined evidence of clinically important potential bias as an interaction-to-main effect ratio of 25% or greater and applied a p-value threshold of 0.10 on the interaction term in the growth curve models. We use a higher p-value criterion than the conventional $p<.05$ because when

evaluating bias, the risk of type 2 error (failing to detect an association) seems more important than the risk of type 1 error (incorrectly inferring an association).

To illustrate the effect of interaction bias, we extracted coefficients from the growth curve models and depicted the trajectory of (a) the predicted neuropsychological test score as a function of the MRI biomarker for low and high education individuals, based on models with or without the MRI biomarker by education interaction and (b) the predicted trajectory from age 75 to 85 if a sudden 1 standard deviation (SD) shift in the MRI biomarker occurs at age 80, for high and low education individuals, again based on models with or without an MRI biomarker by education interaction.

All analyses report two-sided tests and 95% confidence intervals.

Results:

Education was associated with baseline values of MMSE, Z-score and IRT-cognition (Table 1), but not with normalized WMLV (p=0.55), HCV (p=0.63), or CSFV (p=0.26).

In age- and sex-adjusted linear regression models, the associations of WMLV, and CSFV with baseline MMSE were modified by education (Table 2). In each case the interaction met the criterion for clinically relevant bias, such that the association between the MRI measures and MMSE was attenuated or null in people with high levels of education (Table 2, top panel). For example, a 1 SD higher WMLV was associated with 0.17 SD lower MMSE among individuals with low education, but only 0.04 SD lower MMSE for the high education group (p-value for interaction = 0.005).

The main effects of WMLV and HCV on Z-score and IRT-Cognition were significant, and education did not modify these associations. Although the main effects of CSFV on Z-score and IRT-Cognition were not significant, there was evidence that CSFV had a more extreme association with Z-score (interaction p = 0.10) among high education compared to low education individuals; there was no evidence for an interaction in predicting IRT-Cognition (interaction p = 0.26).

In mixed models of longitudinal trajectories of cognitive variables, we found clinically significant education bias in predicting the level of MMSE from WMLV (interaction p = 0.04) and CSFV (interaction p = 0.09) and in predicting the slope of MMSE decline for HCV (interaction p = 0.05). (Figure 2 - blue bars, and eTable 1).

For level of Z-score there was no evidence of bias. For the slope of Z-score decline, there was evidence of clinically significant bias for the CSFV measure only (Figure 2 - orange bars, and eTable 1, interaction p = 0.06). There was no evidence for significant bias for any MRI measure predicting either level or slope of IRT-Cognition (Figure 2 - grey bars, and eTable 1).

To illustrate the impact of education interaction in MMSE, Figure 3a and 3b shows predicted associations of MMSE (Figure 3a) or IRT-Cognition (Figure 3b) as a function of WMLV for high and low education individuals. Higher levels of WMLV are associated with much larger

differences in MMSE for low education individuals than high education individuals, whereas when the IRT-Cognition measure is used, the differences associated with increases in WMLV are nearly parallel. Figure 3c and 3d illustrates a predicted trajectory of MMSE (Figure 3c) and IRT-cognition (Figure 3d) from age 75 to 85 years, assuming that WMLV suddenly increases by 1 SD at age 80. Such an increase might not be detected for high education individuals with the MMSE but would be detectable using the IRT-Cognition measure about equally well regardless of level of education.

Discussion

In this longitudinal, population-based study of 1,782 older French adults, we confirmed that educational level was an important modifier of the link between MRI-based measures of dementia risk and neuropsychological test scores. Every MRI marker predicted MMSE, and for two of three MRI markers this association was modified by education level. This modification implies that the MMSE is a biased marker of underlying disease progression. Bias was generally smaller when evaluating slope (rate of change) in a composite cognitive measure instead of MMSE alone, and bias was clinically and statistically negligible for both level and slope when using an IRT-based measure of cognition as the outcome.

Our study is consistent with prior research showing that higher education masks the association of MMSE with neuroimaging markers associated with dementia risk[3,4] or cerebrovascular disease[5]. The current analyses advances previous work in this area by examining an IRT-based cognition measure and evaluating links with rate of change in each cognitive assessment. We know of only one prior study that has evaluated education interactions between a biological marker of dementia risk and change in cognition in a population sample[6]; this study reported plasma β-amyloid was more strongly associated with declines in Modified Mini-Mental State Examination Scores in low education individuals, consistent with our findings.

A major challenge in interpreting these results is determining whether to conceptualize these interactions as *bias* or as substantive indicators of cognitive reserve. The observed interaction effects of brain variables and education on MMSE scores could reflect measurement bias of the MMSE. The MMSE has been shown to have ceiling effects such that it is less sensitive to individual differences in those with high baseline levels of cognitive test performance compared to those with lower baseline cognitive test performance.[21]. In effect, a 1-point difference between MMSE scores of 29 and 30 represents a much larger difference in actual cognitive ability than does a 1-point difference between scores of 20 and 21. A given brain impact on cognition in those with low and high levels of cognitive function would correspond to a smaller difference in MMSE score for those with higher cognitive function even though the true impact on cognition is the same. Since education is strongly related to MMSE score, this would mean that the same brain difference would result in a smaller MMSE score difference in those with higher education, which is exactly what we observed. The IRT based composite score, in contrast, should not have a different scale of measurement for those with low and high cognitive ability so results for that measure better reflect true brain effects on cognition in the low and high education groups.

Brain by education interactions on cognitive test scores also have been interpreted as evidence that education preserves function in a meaningful sense despite neurologic damage related to AD or other progressive diseases. [2–6] In distinguishing these alternative interpretations, we note that a diagnosis of dementia depends on the *net* impact of multiple factors: cognitive reserve; progressive pathologies adversely affecting the brain, such as amyloid plaques, tau tangles, or cerebrovascular disease; and neurological or functional plasticity facilitating recovery for example in the aftermath of cerebrovascular accidents or the gradual accumulation of brain changes consistent with AD. It is important for researchers to evaluate the determinants of *each* of these three processes: the determinants of cognitive reserve may be quite different from the determinants of progressive brain disease. The appropriate methodological approach differs depending on which of these outcomes is of primary interest.

Many studies which fundamentally aim to shed light on the drivers of brain disease must rely on measures of cognition or dementia because they do not have direct measures of neuropathology. Cognitive reserve may bias the estimates of the effect of a risk factor on brain health because reserve weakens the link between neuropathology and cognition or dementia diagnoses. This motivates the use of education-adjusted norms when interpreting neuropsychological test scores. Using longitudinal analyses has a major advantage over norms, however, because applying educational adjustments to scores makes it harder to detect effects of education on disease.

The goal of this paper was to evaluate whether using longitudinal analyses of IRT-based measurement scales would enable use of neuropsychological assessments as unbiased reflections of at least some MRI based measures of disease.

A critical limitation in this study, as in all cognitive aging research, is the lack of a conclusively verified, uncontested gold standard for measuring cognition. As a result, validation of any cognitive measure is based on indirect approaches, such as demonstrating strong associations with correlates of cognition. An important question for future research is whether change in IRT-based cognitive composites predict outcomes of greatest relevance to patients, such as functional dependence and mortality.

We considered only a few commonly used MRI biomarkers; extension of this work to include other biomarkers strongly related to AD is a critical future direction.[22] The 3C sample does not represent the level of racial/ethnic or linguistic diversity found in many US based samples, but this makes the results even more striking: even in a fairly socially homogeneous sample, the neuropsychological measures were not operating consistently across educational groups. In addition, of the 1,923 people invited to participate in the MRI scan, valid measures could only be obtained from 1,784 participants. Participants who did not contribute valid MRI data were on average older, lower educated, and performed worse on cognitive tests. The analytic sample is therefore likely to represent a healthier subsample of the 3C study participants.

Research on cognitive aging, incidence, and progression of dementia is hampered by the lack of unbiased instruments that can be used to measure general cognitive function independent of the patient's socio-economic status and educational background.

The limitations of common neuropsychological assessments have fostered calls for increasing reliance on biomarker based disease assessments[23], but exclusive reliance on biomarkers is expensive and threatens the diversity of research samples[1], and most biomarker positive people do not go on to develop cognitive impairment or dementia.[24] Our findings indicate that using IRT-based measurement scales and focusing on rates of change, instead of measures based on cross-sectional level of functioning, will enable use of neuropsychological assessments as valid proxies for at least some MRI based measures of disease. Such approaches will improve the validity of research on determinants of cognitive aging and dementia and improve our chances of identifying true causes of disease, rather than only determinants of cognitive reserve or biased test performance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Study funding:

Financial support was from the National Institute on Aging, including AG051170 (SW, ALG, DM, RNJ, TJF, and MMG) and AG056164 (JJM).

The 3C Study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, MGEN, Institut de la Longévité, Conseils Regionaux of Aquitaine and Bourgogne, Fondation de France, and Ministry of Research–INSERM Programme "Cohortes et collections de données biologiques.

Sponsors were not involved in the design of the study or in the data analyses or manuscript elaboration and the authors have no conflicts of interest

References

1. Glymour MM, Brickman AM, Kivimaki M, et al. Will biomarker-based diagnosis of Alzheimer's disease maximize scientific progress? Evaluating proposed diagnostic criteria. Eur J Epidemiol. 2018;33(7):607–612. doi:10.1007/s10654-018-0418-4 [PubMed: 29948371]
2. Bennett DA, Wilson RS, Schneider JA, et al. Education modifies the relation of AD pathology to level of cognitive function in older persons. Neurology. 2003;60(12):1909–1915. [PubMed: 12821732]
3. Dufouil C, Alperovitch A, Tzourio C. Influence of education on the relationship between white matter lesions and cognition. Neurology. 2003;60(5):831–836. [PubMed: 12629242]
4. Perneczky R, Wagenpfeil S, Lunetta KL, et al. Education attenuates the effect of medial temporal lobe atrophy on cognitive function in Alzheimer's disease: the MIRAGE study. J Alzheimer's Dis. 2009;17(4):855–862. [PubMed: 19542606]
5. Elkins JS, Longstreth WT, Manolio TA, Newman AB, Bhadelia RA, Johnston SC. Education and the cognitive decline associated with MRI-defined brain infarct. Neurology. 2006;67(3):435–440. [PubMed: 16894104]
6. Yaffe K, Weston A, Graff-Radford NR, et al. Association of plasma β-amyloid level and cognitive reserve with subsequent cognitive decline. JAMA. 2011;305(3):261–266. [PubMed: 21245181]

7. Cyceron G Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. Neuroepidemiology. 2003;22:316–325. [PubMed: 14598854]
8. Kaffashian S, Tzourio C, Soumaré A, et al. Association of plasma β-amyloid with MRI markers of structural brain aging the 3-City Dijon study. Neurobiol Aging. 2015;36(10):2663–2670. [PubMed: 26242707]
9. Godin O, Tzourio C, Maillard P, Mazoyer B, Dufouil C. Antihypertensive treatment and change in blood pressure are associated with the progression of white matter lesion volumes: the Three-City (3C)-Dijon Magnetic Resonance Imaging Study. Circulation. 2011;123(3):266–273. doi:10.1161/CIRCULATIONAHA.110.961052 [PubMed: 21220733]
10. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189–198. [PubMed: 1202204]
11. Benton A Manuel pour l'application du test de rétention visuelle. Applications cliniques et expérimentales, ed. 1965;2.
12. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. Percept Mot Skills. 1958;8(3):271–276.
13. Lezak M, Metz-Lutz M. Neuropsychological Assessment 3e edition. Neurophysiologie Clinique/Clinical Neurophysiology. 1996;2(26):120–121.
14. Isaacs B, Kennie AT. The Set test as an aid to the detection of dementia in old people. Br J Psychiatry. 1973;123(575):467–470. [PubMed: 4748864]
15. Rosseel Y lavaan: An R package for structural equation modeling. Journal of Statistical Software. 2012;48(2):1–36.
16. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2015.
17. Kamata A, Bauer DJ. A note on the relation between factor analytic and item response theory models. Structural Equation Modeling. 2008;15(1):136–153.
18. Gross AL, Sherva R, Mukherjee S, et al. Calibrating longitudinal cognition in Alzheimer's disease across diverse test batteries and datasets. Neuroepidemiology. 2014;43(3–4):194–205. doi:10.1159/000367970 [PubMed: 25402421]
19. Gross AL, Power MC, Albert MS, et al. Application of Latent Variable Methods to the Study of Cognitive Decline When Tests Change over Time. Epidemiology. 2015;26(6):878–887. doi:10.1097/EDE.0000000000000379 [PubMed: 26414855]
20. Brown TA. Confirmatory Factor Analysis for Applied Research. Guilford Publications; 2015.
21. Mungas D, Reed BR. Application of item response theory for development of a global functioning measure of dementia with linear measurement properties. Stat Med. 2000;19(11–12):1631–1644. [PubMed: 10844724]
22. Lockhart SN, DeCarli C. Structural imaging measures of brain aging. Neuropsychol Rev. 2014;24(3):271–289. [PubMed: 25146995]
23. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia. 2018;14(4):535–562. doi:10.1016/j.jalz.2018.02.018
24. Brookmeyer R, Abdalla N, Kawas CH, Corrada MM. Forecasting the prevalence of preclinical and clinical Alzheimer's disease in the United States. Alzheimers Dement. 2018;14(2):121–129. doi:10.1016/j.jalz.2017.10.009 [PubMed: 29233480]

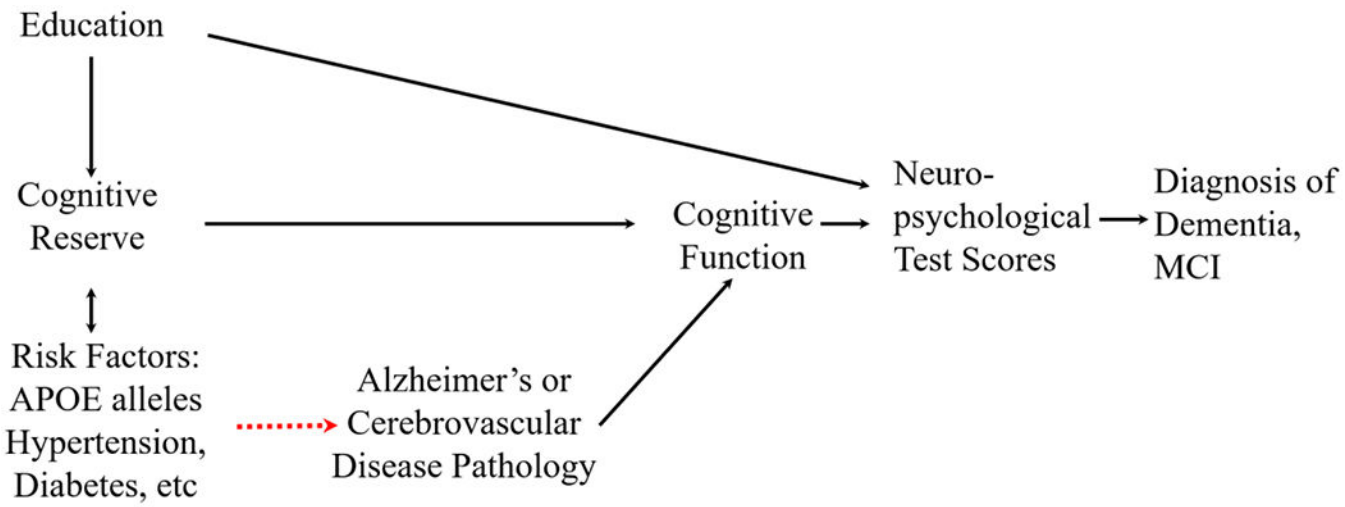

Figure 1. Conceptual model linking education, neuropsychological test scores, and dementia diagnoses.

Neuropsychological measures will be biased as tools for evaluating determinants of underlying disease if differences in neuropsychological test performance do not equate to the same differences in the disease for high or low education individuals. In such a setting, education influences neuropsychological test scores independently of underlying disease, so any correlates of education will predict neuropsychological test scores even if they have no influence on underlying disease pathology. The same biases will apply for any diagnostic outcomes that rely on neuropsychological test scores.

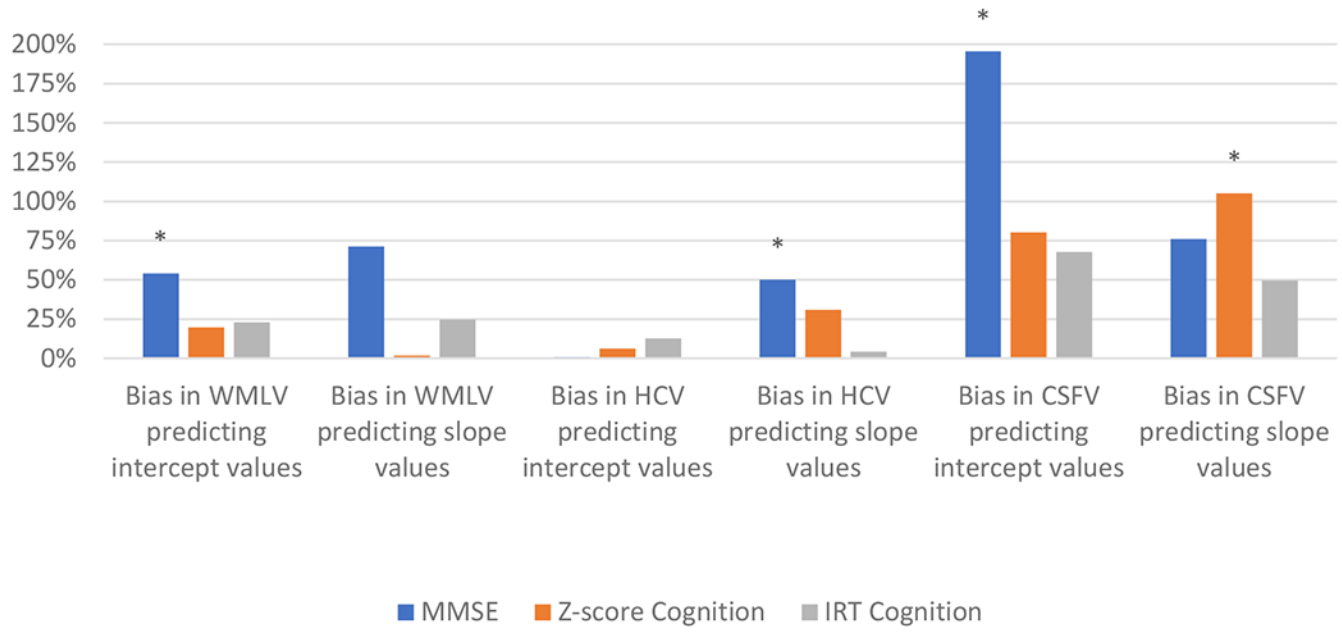

Figure 2: Percent bias of estimated association of neuroimaging markers with intercept or slope of MMSE, Z-score composite, and IRT-Cognition by education level.

Bias is expressed as the ratio of the interaction coefficient to the main effect coefficient of the neuroimaging marker on each outcome. Estimates are from linear mixed models with age and sex adjustment and random intercepts and slopes (regression model coefficients shown in Appendix). Estimates that meet the criteria for clinical significance (greater than 25% and based on interactions with p-value<0.10) are denoted with an *. Note that as a ratio measure, the percent bias may be large if either the interaction (the numerator) is large or the main effect of the MRI measure on the neuropsychological assessment (the denominator) is very small.

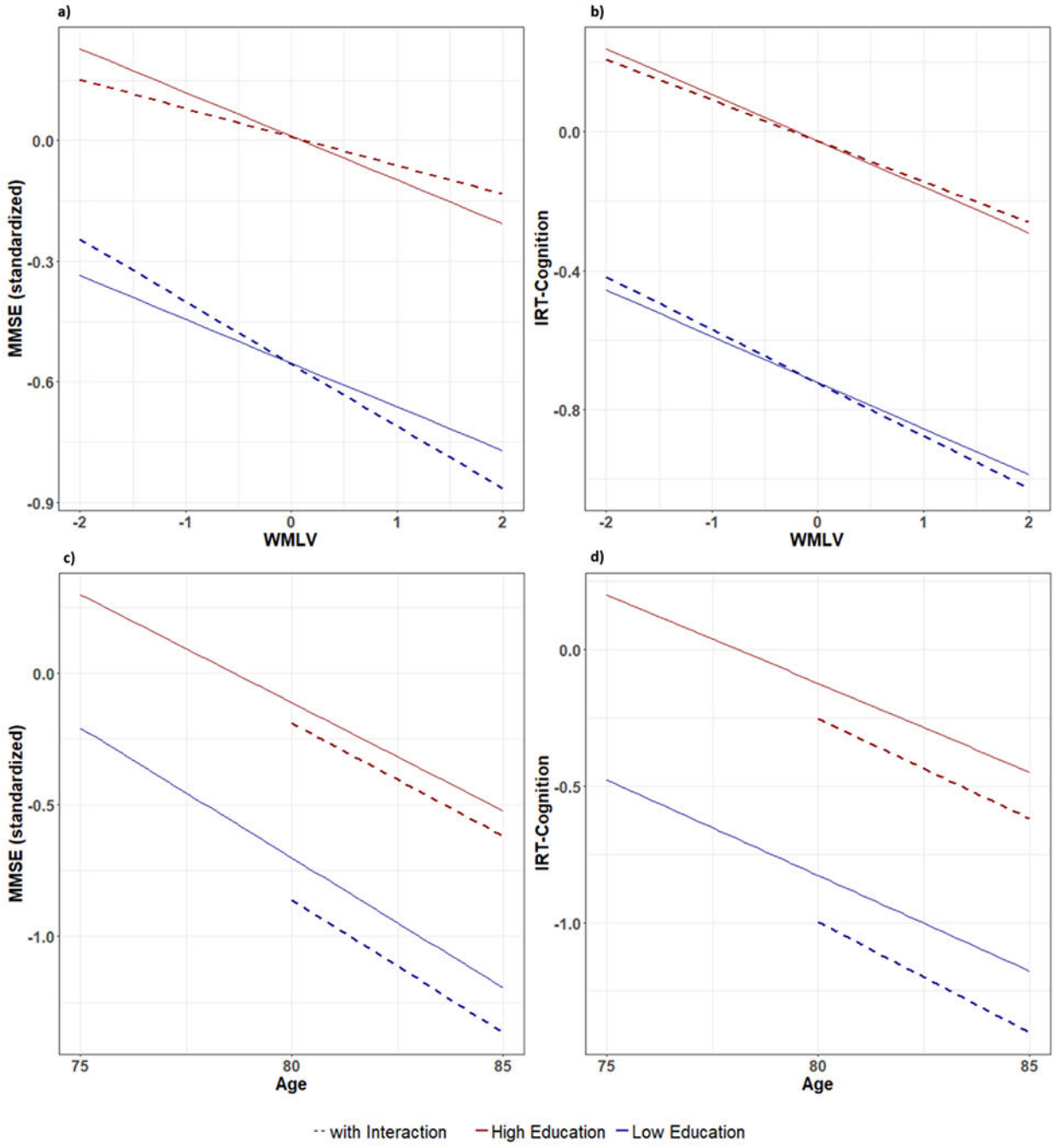

Figure 3: Comparing predicted cognitive outcomes for low and high education individuals based on models allowing an education interaction versus models with no interaction.

Models with no interaction assume an identical effect of MRI differences on cognition, regardless of education. Figures show predictions for: (a) MMSE as a function of WMLV and education, and (b) IRT-Cognition as a function of WMLV and education. For MMSE, omitting the education interaction substantially over-estimates WMLV-related change for high education individuals but under-estimates WMLV-related change for low education individuals; thus, models assuming the association between WMLV and MMSE is similar regardless of education are severely misspecified. The misspecification is smaller and non-

significant in the models using IRT-Cognition. When comparing predicted cognitive measures before and after a sudden increase in WMLVplots show predictions for low and high education individuals who experience typical age related changes followed by a sudden 1 SD increase in WMLV at age 80, showing predicted values for (a) MMSE; and (b) IRT-Cognition. The sudden increase in WMLV would have little impact on the MMSE of high education individuals, but would lead to a sudden decline in MMSE in the low education individuals. The sudden increase in WMLV would have a similar impact on IRT-Cognition for high and low education individuals. This means that the MMSE does not serve to identify a sudden increase in WMLV on high education individuals, but the IRT-Cognition measures does.

Table 1.

Baseline characteristics of participants in the 3 City Dijon study neuroimaging sample, stratified by education.

	All	Low education (≤8 years of school)	High education (>8 years of school)	Hedges's g calculation for association with education[1]	P-value test for the null of no difference by education
n	1782	818 (46%)	964 (54%)		
Age (mean, SD)	72.40 (4.10)	72.48 (4.02)	72.33 (4.17)		0.447
Sex (n, %)	1075 (60%)	526 (64%)	549 (57%)		0.002
Mini Mental State Examination, raw (mean, SD)	27.65 (1.82)	27.14 (2.02)	28.08 (1.51)	−0.62	<0.001
Mini Mental State Examination, standardized (mean, SD)	0.00 (1.00)	−0.28 (1.11)	0.24 (0.83)	−0.62	<0.001
Z-Score Cognitive Composite (mean, SD)	0.00 (1.00)	−0.35 (0.98)	0.30 (0.92)	−0.71	<0.001
IRT-Cognition (mean, SD)	0.00 (1.00)	−0.35 (0.98)	0.30 (0.91)	−0.74	<0.001
Intracranial Volume, raw (mean, SD)	1364.74 (136.03)	1345.25 (137.24)	1381.29 (132.83)	−0.27	<0.001
White Matter Lesion Volume cm^3, raw (mean, SD)	5.43(4.38)	5.39 (4.31)	5.47 (4.45)	−0.02	0.687
White Matter Lesion Volume, normalized (mean, SD)	0.00 (1.00)	0.02 (0.99)	−0.01 (1.01)	+0.03	0.555
Hippocampal Volume, raw (mean, SD)	6.62 (0.83)	6.54 (0.83)	6.69 (0.82)	−0.18	<0.001
Hippocampal Volume, normalized (mean, SD)	0.00 (1.00)	−0.01 (1.01)	0.01 (0.99)	−0.02	0.627
Cerebrospinal Fluid Volume, raw (mean, SD)	385.86 (62.69)	380.29 (62.92)	390.58 (62.14)	−0.17	0.001
Cerebrospinal Fluid Volume, normalized (mean, SD)	0.00 (1.00)	0.03 (0.99)	−0.02 (1.01)	+0.05	0.261

[1] Hedges's g calculated as mean in the high education group minus mean in the low education group, divided by the pooled weighted standard deviation.

Alzheimer Dis Assoc Disord. Author manuscript; available in PMC 2021 February 25.

Alzheimer Dis Assoc Disord. Author manuscript; available in PMC 2021 February 25.

Table 2

Linear regression coefficients for the association between brain MRI markers and baseline neuropsychological outcomes (MMSE, Z-score composite, or IRT-Cognition, measured in 1/10 of SD), estimated including interaction terms between education and MRI markers. *

	White Matter Lesion Volume		Hippocampal Volume		Cerebral Spinal Fluid Volume	
	Beta (95% CI)	**P-value**	**Beta (95% CI)**	**P-value**	**Beta (95% CI)**	**P-value**
Predicting MMSE						
MRI Measure	−1.68 (−2.35,−1.02)	<0.001	−0.21 (−0.9,0.48)	0.547	1.11 (0.41,1.81)	0.002
Education	5.13 (4.24,6.03)	<0.001	5.11 (4.21,6.01)	<0.001	5.13 (4.23,6.03)	<0.001
Education * MRI Measure	1.27 (0.37,2.16)	0.005	0.7 (−0.2,1.59)	0.128	−1.23 (−2.13,−0.33)	0.007
Predicting Z-Score Cognition						
MRI Measure	−1.14 (−1.77,−0.5)	<0.001	0.75 (0.09,1.41)	0.026	0.16 (−0.51,0.84)	0.633
Education	6.51 (5.65,7.38)	<0.001	6.51 (5.65,7.37)	<0.001	6.51 (5.64,7.37)	<0.001
Education * MRI Measure	0.49 (−0.37,1.35)	0.262	0.21 (−0.64,1.07)	0.628	−0.73 (−1.6,0.13)	0.095
Predicting IRT-Cognition						
MRI Measure	−1.01 (−1.64,−0.38)	0.002	0.79 (0.14,1.45)	0.017	−0.12 (−0.79,0.54)	0.718
Education	6.72 (5.87,7.57)	<0.001	6.71 (5.86,7.56)	<0.001	6.7 (5.85,7.56)	<0.001
Education * MRI Measure	0.26 (−0.59,1.11)	0.545	0.34 (−0.51,1.19)	0.431	−0.49 (−1.34,0.37)	0.263

* Education is coded as a binary variable indicating whether the individual completed more than 8 years of schooling.