



HHS Public Access

Author manuscript

Alzheimers Dement. Author manuscript; available in PMC 2015 November 22.

Published in final edited form as:

Alzheimers Dement. 2015 September ; 11(9): 1098–1109. doi:10.1016/j.jalz.2015.06.1885.

Guidelines for reporting methodological challenges and evaluating potential bias in dementia research

Jennifer Weuve^{a,1}, Cécile Proust-Lima^{b,1}, Melinda C. Power^{c,1}, Alden L. Gross^{c,d}, Scott M. Hofer^e, Rodolphe Thiébaud^{b,f,g}, Geneviève Chêne^{b,f,g}, M. Maria Glymour^{h,i,2}, and Carole Dufouil^{b,f,g,*,2} for the MELODEM Initiative³

^aRush Institute for Healthy Aging, Rush University Medical Center, Chicago, IL, USA

^bINSERM U897, Epidemiology and Biostatistics Center, Bordeaux School of Public Health, Bordeaux University, Bordeaux, France

^cDepartment of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD, USA

^dJohns Hopkins Center on Aging and Health, Baltimore, MD, USA

^eDepartment of Psychology and Centre on Aging, University of Victoria, Victoria, BC, Canada

^fClinical Investigation Center-Clinical Epidemiology-CIC-1401 of INSERM U897, Bordeaux, France

^gBordeaux University Hospital (Public Health Department), Bordeaux, France

^hDepartment of Social and Behavioral Sciences, Harvard School of Public Health, Boston, MA, USA

ⁱDepartment of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA

Abstract

Clinical and population research on dementia and related neurologic conditions, including Alzheimer's disease, faces several unique methodological challenges. Progress to identify preventive and therapeutic strategies rests on valid and rigorous analytic approaches, but the

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author. Tel.: +3-355-757-1423; Fax: +3-355-757-5713. carole.dufouil@inserm.fr.

¹Have contributed equally as first authors.

²Have contributed equally as last authors.

³MELODEM initiative (in last name alphabetical order): Amouyel Philippe, Andrieu Sandrine, Beiser Alexa S, Bellenguez Céline, Benitez Andreana, Bennett David, Bouteloup Vincent, Bottai Matteo, Brayne Carol, Chang Chung-Chou, Chibnik Lori, Chupin Marie, Commenges Daniel, De-Carli Charles, Dodge Hiroko, Farewell Daniel, Fratiglioni Laura, Ganguli Mary, Haan Mary, Ikram M Arfan, Jacqmin-Gadda Hélène, Jessen Frank, Joly Pierre, Keiding Niels, Launer Lenore J, Leffondré Karen, Maillard Pauline, Mangin Jean-François, Manly Jennifer, Marioni Riccardo, Mayeda Elizabeth Rose, Mirza Saira, Mosley Thomas, Muniz Terrera Graciela, Pankratz Shane, Pernecky Robert, Petersen Ronald, Piccinin Andrea, Santoni Giola, Seshadri Sudha, Skoog Ingmar, Stern Yaakov, Tchetgen Tchetgen Eric, Thorvaldsson Valgeir, Vivot Alexandre, Xu Weili, Yaffe Kristine, and Zahodne Laura.

Disclosures: C.D. has received personal compensation for teaching activities with American Academy of Neurology. The other authors have no disclosures.

Supplementary data

Supplementary data related to this article can be found online at <http://dx.doi.org/10.1016/j.jalz.2015.06.1885>.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. None of the funding source has any influence on the content of this position article.

research literature reflects little consensus on “best practices.” We present findings from a large scientific working group on research methods for clinical and population studies of dementia, which identified five categories of methodological challenges as follows: (1) attrition/sample selection, including selective survival; (2) measurement, including uncertainty in diagnostic criteria, measurement error in neuropsychological assessments, and practice or retest effects; (3) specification of longitudinal models when participants are followed for months, years, or even decades; (4) time-varying measurements; and (5) high-dimensional data. We explain why each challenge is important in dementia research and how it could compromise the translation of research findings into effective prevention or care strategies. We advance a checklist of potential sources of bias that should be routinely addressed when reporting dementia research.

Keywords

Alzheimer disease; Dementia; Neuropsychological tests; Longitudinal studies; Epidemiologic factors; Statistical models; Selection bias; Survival bias; Big data; Genomics; Brain imaging

1. Introduction

Despite more than two decades of research on prevention and treatment of dementia and aging-related cognitive decline, highly effective preventive and therapeutic strategies remain elusive. Many features of dementia render it especially challenging: supposedly distinct underlying pathologies lead to similar clinical manifestations, development of disease occurs insidiously over the course of years or decades, and the causes of disease and determinants of its severity are likely multifactorial. However, progress in preventing and treating dementia also rests on how dementia research is conducted: informative research requires valid and rigorous analytic approaches, and yet the research literature reflects little consensus on “best practices.”

Several methodological challenges arise in studies of the determinants of dementia risk and cognitive decline. Some challenges, such as unmeasured confounding or missing data, are common in many research areas; others, such as outcome measurement error and lack of a “gold standard” outcome assessment, are more pervasive or more severe in dementia research [1–3]. Currently, researchers handle these challenges differently, making it difficult to directly compare studies and combine evidence. Although some methodological differences across studies arise because analytic methods are explicitly tailored to the study design and realities of the data at hand, other differences arise for less substantive reasons. Modifiable sources of inconsistency include the absence of consensus and definitive standards for best analytic approaches; different disciplinary traditions in epidemiology, clinical research, biostatistics, neuropsychology, psychiatry, geriatrics, and neurology; and software and technical barriers.

The various analytic methods used in dementia research often address subtly distinct scientific questions, depend on different assumptions, and provide differing levels of statistical precision. Unfortunately, there is often insufficient attention to whether a chosen method addresses the most relevant scientific question and relies on plausible assumptions. Some common methods likely provide biased answers—i.e., answers that diverge

systematically from the truth—to the most relevant scientific questions. Even if several alternative approaches might be appropriate and innovative or novel analyses used in individual studies may be valuable, it can be advantageous to report results using a shared approach [4,5]. The “inconsistent application of optimal methods” within and across studies makes it difficult to qualitatively or quantitatively summarize results across studies (meta-analyses). By contrast, a core set of shared analytic approaches would enhance opportunities to synthesize results and more conclusively address our research questions. Applying a set of standardized sensitivity analyses would help evaluate the plausible magnitude of various sources of bias or violations of assumptions. In randomized clinical trials (RCTs), for example, there are strict rules regarding intention-to-treat analyses, which are often complemented with additional approaches, such as per protocol analysis or modeling the complier average causal effect to account for noncompliance.

The CONSolidated Standards of Reporting Trials (CONSORT) [6] and Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines [7] provide helpful indications of broad relevance in human subjects research but are too broad to address several specific methodological difficulties in dementia research. Topic-specific guidelines building on STROBE have proven useful in several domains, such as genetic association studies [8]. The MEthods in LOngitudinal research on DEMentia (MELODEM) initiative was formed in 2012 to address these difficulties and achieve greater consistency in the process of selecting and applying preferred analytic methods across research on dementia risk and cognitive aging. The initial MELODEM findings outline a set of methodological problems that should routinely be addressed in dementia research, summarized in the guidelines in Fig. 1. We advance this list as a working set of guidelines for transparent reporting of methods and results and therefore the best chance of accelerating scientific progress in identifying determinants as well as validating biomarkers for earlier diagnosis of Alzheimer’s disease (AD). The goals of MELODEM include fostering methodological innovation to address these challenges and improving understanding of tools to address each challenge. In this initial report from MELODEM, we focus on outlining major categories of bias and why they are especially relevant in dementia research. We briefly discuss in the following, with more details in the online supplement, five major challenges: (1) selection, i.e., handling selection stemming from study participation, attrition, and mortality; (2) measurement, i.e., dealing with the quality of measurements of exposure and outcomes and how imperfect measurement quality affects analysis and interpretation of results; (3) alternative timescale, i.e., specifying the time-scale and the shape of trajectories in longitudinal models; (4) time-varying exposures and confounding, i.e., accounting for changes in explanatory variables; and (5) high-dimensional data, i.e., analyzing complex and multidimensional data such as neuroimaging, genomic information, or database linkages.

For some topics in the checklist (Fig. 1), substantial controversy remains regarding optimal analytic approaches, especially when considering both bias and variance of the methods. In many cases, although the potential for bias is clear, it has not been established that this bias is substantial in real data. The guidelines in Fig. 1 are intended as a first step toward improved evaluation and reporting of methodological challenges in dementia research, to

support a move toward field-wide consensus on best practices, and identifying the highest priority areas for methodological innovations.

2. Major methodological challenges in longitudinal research on dementia

2.1. Defining the scientific question of interest

Epidemiologic research in cognitive aging identifies correlates of cognitive performance, cognitive decline, and dementia. A critical first step in conducting effective research is often taken for granted: clearly defining the scientific question and distinguishing whether this question is causal or predictive [9]. Research on the etiology of dementia or evaluating potential preventive or therapeutic interventions seeks to address causal questions. Predictive studies, designed for estimating prevalence or identifying high-risk individuals, are important for anticipating future trends in public health burden of dementia and individual patient outcomes. The analytic concerns of predictive studies differ from those of studies intended to support causal inferences.

Observational research is often critiqued for drawing unsubstantiated causal inferences, but providing evidence to support causal inference is commonly the primary goal of statistical analysis. Although conclusive inferences for many treatments may ultimately depend on evidence from RCTs, intervention trials are typically fielded only after extensive observational research. This observational evidence should be accumulated with the goal of illuminating causal structures and guiding design of (eventual) trials.

2.2. Selection into the analysis sample

Selection issues in dementia research arise from differential attrition of enrolled participants, differential survival of enrolled participants, and differential enrollment, either due to refusal to participate or differential survival up to the moment of study initiation (Table 1 and Supplementary Appendix 1). Each of these processes can bias effect estimates. Spurious associations between the putative risk factor and cognitive decline or dementia can occur when selection processes are related to cognitive status and the exposure of interest (or their determinants). The bias is not necessarily toward the null (i.e., which would tend to mask an association) and can sometimes reverse the direction of association (i.e., making harmful exposures appear protective or protective exposures appear harmful) [10]. Impaired cognition and dementia have broadly debilitating consequences, setting the stage for selection bias in longitudinal studies of dementia; risk of illness [11–15], death [16,17], and study attrition [18–20] are all heightened among persons with impaired cognition. Selection bias is likely to result if the risk factor under study is associated with attrition as well. For example, chronic disease, adverse health behaviors, socioeconomic status, and race are all associated with substantial morbidity and mortality risks [3,21–25]. Various approaches are adopted to address potential selection bias (Table 1), but the performance of these approaches has rarely been evaluated. Even the likelihood of substantial bias from “ignoring the selection process,” which is arguably the most common strategy in the applied literature, is rarely formally quantified, although specific examples suggest it may be reasonably large. In the Chicago Health and Aging Project, accounting for selective attrition increased estimated associations between smoking and cognitive decline by 56%–86% [26].

MELODEM guidelines are intended to provide evidence to evaluate whether selection is likely to introduce a substantial bias.

Over the course of longitudinal follow-up, many older participants may “drop out” or refuse continued study participation. If drop-out is dependent on measured parameters (a “missing at random” mechanism), several analytical approaches can provide unbiased effect estimates; however, if drop-out is dependent on unknown or unmeasured parameters, there is no easy solution for bias correction (Table 1). In this situation, sensitivity analyses can illuminate the robustness of the findings [27].

Mortality is also a significant source of censoring in longitudinal studies. Many approaches considered appropriate for handling dropout-related attrition are more controversial in the context of survival, and there is no current consensus on preferred approaches [28]. Although cognitive function is “missing” for individuals who drop out of a study, it is more appropriately described as “undefined” for people who die [29]. This conceptualization suggests that people whose survival is determined by exposure should be excluded from the population for whom we try to estimate effects. For example, smoking may predict lower dementia diagnosis rates by causing earlier mortality from other causes [2]. Thus, the central challenge when addressing selective survival is to clearly define the question of scientific interest [28,29]: what parameter are we trying to estimate, for whom are we trying to estimate it, and which analysis methods correspond with this estimand (Table 1)?

In dementia cohort studies, truncation of follow-up by death also introduces interval censoring, which occurs because diagnosis of dementia can only be made at periodic follow-up visits. Therefore, dementia status at death is unknown for participants who were free of dementia at their last visit before death. Interval censoring in the presence of competing risk of death can induce an underestimation of dementia incidence and alter estimated effects of exposures [23,25]. For example, the protective effect of high education on the risk of dementia was overestimated by 36% in men when not accounting for interval censoring in the French PAQUID cohort. This was likely because, while higher educational attainment predicts elevated risk of dementia diagnosis, it also predicts faster death *after* the diagnosis [25].

Bias may also arise from how participants are selected into a study if enrollment is influenced by the exposure of interest and the outcome (or their determinants). Similar biases may occur whether differential enrollment occurs because people with particular dementia risk factors systematically refuse study participation or because these people are unlikely to survive to the age of enrollment. Consider a hypothetical study of the effect of smoking on AD, enrolling participants at age 70 (Fig. 2). If the effects of smoking and *APOE-ε4* status on mortality are synergistic (i.e., more than multiplicative), then the surviving 70-year-old smokers will have a lower *APOE-ε4* prevalence than 70-year-old nonsmokers. The study may conflate effects of *APOE-ε4* with the effects of smoking. Pre-enrollment selective mortality is particularly likely in population-based studies of older adults: persons exposed to detrimental risk factors may have survived to the age of enrollment only by virtue of their unusually effective detoxification genotype or cognitive acumen. Challenges arising from selective enrollment are broadly recognized, although the

potential for this selection to compromise both generalizability and internal validity is sometimes disregarded.

Each of these selection processes may partially “mute” effect estimates. For example, the association between smoking and AD progressively attenuates, or even becomes protective, in older samples [2]. Similar muting effects with increased participant age are apparent for multiple other risk factors [2,22,30–33]. Selection may also create spurious differences in effect estimates between subgroups.

Table 1 summarizes these problems and commonly adopted analytic approaches to addressing issues of selection. Evaluating the usefulness of each approach and how each of these approaches perform in specific situations in dementia research is an important methodological question.

Although issues of selective attrition and selective survival affect both observational research and RCTs, selective enrollment does not compromise the internal validity of RCTs (although may reduce generalizability) conferred through randomization.

2.3. Measurement validity and reliability

Measurement challenges in dementia research result from the disjuncture between disease pathophysiology and clinical and research measures, due to imperfect validity and reliability (detailed in Supplementary Appendix 2). For example, performance on neuropsychological tests does not necessarily precisely reflect biological functioning and capacity of the brain. Similar measurement challenges pertain to measures of the consequences, severity, and progression of dementia, including functional dependency, neuropsychiatric symptoms, and behavioral patterns [34]. Distinguishing between neuropathologic processes in the brain and cognitive symptom trajectories is important to elucidate specific causal pathways to disease and possible interactive effects on closely related outcomes.

Validity refers to whether a measurement instrument assesses the phenomenon of interest. Gold standard measures can be used to assess the validity of alternative measurements, but there is often no clear gold standard in dementia research. Measures valid for one group of people may not be valid for another, leading to biased estimates of disparities and risk factor effects [35,36]. For example, historical inequalities in educational access for African-Americans compared with white Americans have led to systematic differences in literacy levels in older adults. These literacy differences appear to contribute to racial disparities in dementia risk [35,36].

Identifying valid dementia biomarkers is critical but many efforts risk circular reasoning, in which we use clinical diagnoses to validate biomarkers and those same biomarkers to validate the clinical diagnoses [37]. Indeed, even the phenotype of interest is often controversial, and it is likely that many common disease definitions include diverse underlying pathologies.

Reliability is the proportion of variability in a measure explained by the construct of interest, as opposed to the proportion attributable to measurement error (random fluctuations in the measurement not reflecting changes in the underlying construct) [38,39]. Nearly all

neuropsychological assessments have substantial unreliability, which reduces statistical power and introduce the potential for regression to the mean.

Practice or retest effects arising from changes in familiarity with the testing process, use of strategies, or recall of test-specific content can also hamper detection of cognitive decline [40] because practice effects can be large enough to offset several years of cognitive decline in elderly adults [41]. Cognitive declines stemming from incipient dementia may thus be impossible to detect because of practice-related improvements on test performance [41]. Practice effects could lead to underestimation of the rate of cognitive decline, incidence of dementia, and, if the magnitude of practice effects differs by background characteristics, to incorrect inferences about determinants of cognitive decline [41,42]. Statistical approaches for handling practice effects remain controversial (Table 2) [43,44].

Another measurement challenge comes from the assumption made by most common analytic methods that a 1-point difference in a test score has the same substantive meaning at high and low ends of the scale. A decline in Mini-Mental State Examination (MMSE) score from 25 to 24, however, may not be equivalent to a decline from 20 to 19. Ceilings and floors constitute extreme examples of such unequal interval scaling. Ceilings and floors attenuate effect estimates in cross-sectional analyses but may either attenuate or inflate effect estimates in longitudinal analyses [1]. An extensive simulation study showed that failing to account for unequal interval scaling of psychometric tests when studying effects of a risk factor on cognitive slope can substantially inflate type 1 errors (i.e., spurious associations) if the risk factor also predicts baseline cognitive level [45].

2.4. Alternative timescales and specification of longitudinal models

The specification of the timescale(s) and functions of within-person change in longitudinal studies can dramatically influence results and replicability. Because of the close link between age and dementia, age constitutes a natural and appropriate timescale for studying dementia risk or related binary outcomes [46–48]. In studies of cognitive ageing, this also applies when studying time-invariant exposures (such as gender or genes). However, research often addresses time-varying exposures that are measured only once during the study (e.g., nutrition, diabetes, treatment) and thus at different ages. In this situation, using the time since exposure measurement (usually enrollment) as the timescale may be more appropriate. When focusing on specific phases of cognitive aging such as the prodromal phase of dementia or terminal decline, reverse time (e.g., years before diagnosis or death) may also be informative [44,49]. However, using reverse time inherently selects participants who developed the outcome, which might cause biases in estimated longitudinal changes. This is an active area of methods development, with several approaches used in the current literature (see section 2.2). Generally, the fundamental underlying causal process presumed to be relevant should guide the choice of metrics [50] (Table 3).

Most observational studies on cognition recruit participants over a wide age range so that studying cognitive change with age mixes two processes: within-person change with age (usually of main interest) and between-person age differences that are also influenced by birth cohort. Ignoring age differences at baseline when studying cognitive decline with age is appropriate only if individuals “converge” onto the same age trajectory whatever their

birth cohort: i.e., if a person entering at 85 years is expected to have the same cognitive level as a person entering at 65 years and followed for 20 years. This may be unrealistic [50,51]. For example, for women in the Whitehall cohorts, between-person age effects overestimated rate of cognitive decline compared with within-person effects because of large cohort differences in educational levels [52]. The “convergence” issue can be easily disentangled by distinguishing two timescales: a longitudinal timescale (e.g., current age or time since enrollment) for within-person change and a cross-sectional timescale (e.g., age at enrollment) for between-person age differences (Table 3).

Average year-to-year cognitive changes are not expected to be the same at all ages; cognitive decline may accelerate at older ages. In studies with short follow-ups, linear approximations may be adequate [53] but with longer follow-up, age-heterogeneous samples, or pathologic events, linearity rarely holds. Approaches to account for this heterogeneity include polynomial cognitive trajectories [54]; biphasic trajectories with change points [55,56] or nonparametric estimation of cognitive trajectories [49].

The Supplementary Appendix 3 provides further detail on the problems of timescales.

2.5. Time-varying exposure/time-varying confounding

Pathologic brain changes are evident at least two decades before clinical dementia diagnosis. Effects of exposure on cognitive outcomes may depend on when exposure occurs, and the relevant timing likely differs for exposures influencing pathogenesis, disease progression, and/or maintenance of function. Research identifying relevant etiologic periods is essential for guiding clinical decisions and preventive interventions targeting known risk factors. For example, elevated blood pressure in midlife predicts higher dementia risk, whereas elevated blood pressure late in life does not [33,57–59]. Although the explanation for this difference is unclear, recommendations on hypertension treatment for dementia prevention must be tailored to a person’s age and existing morbidities. A “critical window” hypothesis has been suggested for hormone therapy effects on dementia, with benefits from initiation in the perimenopausal period but harms from later initiation [60]. A recent systematic review identified only a handful of studies directly addressing this question [61]. For evaluating etiologic periods, cohorts with very long follow-up (e.g., PREVENT [62] or the Framingham Heart Study [63]) are informative because they provide measures of exposure at multiple ages. Quantifying how an exposure’s effects on dementia evolve with age of exposure requires care because different methods of analyzing time-varying exposure data can yield results that vary substantially in magnitude and, more crucially, in interpretation (Table 4) [64]. Although risk factor-dementia associations that differ by age at exposure may reflect relevant etiologic windows, they could also reflect reverse causation or measurement error that differs across age-specific exposures. When cumulative effects of long-term exposure, rather than variation in exposure or point in time exposure effects, are hypothesized, repeated measures of exposure can be combined (e.g., averaged) to achieve more precise exposure estimates.

Just as exposures may vary over time, so too may confounders. Adding to the challenge are differences across studies in how and when exposures and potential confounders are measured. Similar to analyzing time-varying exposure data, adequately adjusting for

confounding when exposures and confounders change over time requires special care, especially when confounders also act as potential mediators [65].

The Supplementary Appendix 4 provides further detail on and examples of the problems of time-varying exposures and confounders.

2.6. High-dimensional data

The proliferation of data sources and emergence of high-dimensional data could powerfully accelerate dementia research but only if harnessed effectively [66]. Administrative databases, omics data, brain imaging data (magnetic resonance imaging and positron emission tomography), biomarker panels assessing gene expression, and metabolic pathways, among many others, present both new challenges and opportunities (Supplementary Appendix 5).

New technologies will provide information on numerous biomarkers, which may help us distinguish more specific dementia phenotypes, but we need strong measurement tools to better take advantage of these data.

Challenges in the statistical analyses of high-dimensional data are numerous (Table 5), even when the number of observations is much larger than typically available in research cohorts. Big data do not necessarily resolve the familiar internal and external validity challenges in epidemiologic studies and may in some situations exacerbate challenges with measurement validity and selection bias. For example, dementia is known to be substantially underrepresented in many US administrative databases, due to underdiagnosis. Underdiagnosis rates may differ by demographic or other background variables [67]. A provocative and unexpected finding was recently reported from a UK administrative database (UK Clinical Practice Research Datalink) with almost 2 million individuals ages 40 years, accruing over 45,000 incident dementia cases [68]: obese people carried a third lower risk of developing dementia than their normal weight peers. The huge sample provided very precise effect estimates but also entailed numerous tradeoffs that may have introduced biases, such as limited confounder control based on variables recorded in the database and potential for misclassification of the outcome [69].

Recent worldwide efforts to merge data from multiple dementia cohorts are generating unique and promising databases: the European Medical Information Framework-AD, a partnership of academics, pharmaceutical companies, and medical informatics specialists focuses on the identification of preclinical biomarkers of AD. Despite huge sample sizes, results will be useful only if it is possible to derive harmonized measures of exposure, biomarkers, and outcomes. Coordinating such efforts is a major challenge [66]. The Integrated Analysis of Longitudinal Studies of Aging collaboration focuses on data harmonization and reproducibility of research from international longitudinal studies focused on aging and health-related change in cognition, health, and well-being [70].

High-dimensional data analysis generates multiple statistical comparisons/tests potentially addressed by various statistical corrections (family-wise error and false discovery). A related issue is possible “overfitting” (i.e. when a statistical model describes random error or noise

instead of the underlying relationship). In this case, dimension reduction techniques can be used either during or before supervised analyses (e.g., Lasso or partial least-squares methods). The challenges here are both to account for the high dimensionality in models (multiple testing and overfitting) and to extract meaningful information from these data to better understand dementia etiology.

New data sources may allow more powerful hypothesis-driven research, i.e., evaluating prespecified social, behavioral, or clinical determinants of dementia, and also agnostic search conducted in the absence of clearly specified hypotheses. Agnostic search approaches are challenging because of type 1 error problems, but this is an important frontier for the field.

With the growing availability of biomarkers in large data sets, there may be new opportunities to evaluate interactive effects of risk factors and pathologic processes. For example, researchers may be able to model whether genetic or behavioral risk factors influence cognitive symptoms differently depending on level of underlying neuropathology. Importantly, because many but not all risk factors will directly influence neuropathology, this will entail careful mediation models to decompose direct, indirect, and interactive pathways.

The Supplementary Appendix 5 provides further detail on the problems of high-dimensional data.

3. Conclusion

Clinical and epidemiologic research over the past two decades has witnessed a remarkable movement toward improved transparency, reproducibility, and methodological rigor, as reflected in the CONSORT [6,71], STROBE [7], and recently STARDdem [37] guidelines. Although the STROBE recommendations apply to studies of dementia, they provide quite general suggestions, with little guidance on many important issues of special salience in dementia and cognitive ageing research. The MELODEM guidelines fill this gap. Embracing and expanding on the foundation of STROBE and related efforts could strengthen and accelerate dementia research.

The MELODEM guidelines, highlighting a set of common methodological challenges in longitudinal research on dementia, complements STROBE and CONSORT guidelines but focuses on technical challenges specific to dementia-related research. We hope researchers adopting the MELODEM guidelines will routinely acknowledge these challenges and justify analytic decisions. We anticipate that MELODEM will provide a platform for continued discussion and innovation of methodological tools to strengthen dementia research.

The guidelines indicate standards for reporting but do not make specific recommendations for how to best address analytic challenges. Ideally, consensus recommendations for “common denominator” analyses will emerge in coming years to facilitate meta-analyses and integration of evidence. Findings from common denominator analyses could routinely be included in research articles, alongside additional methodological approaches based on research team innovations or particular strengths of the data set.

Observational studies suggest major differences in prevalence and incidence of dementia-related phenotypes across population groups; these differences represent opportunities to prevent dementia, if we can identify the reasons for epidemiologic patterns. Such research will only be effective if we can overcome the methodological challenges discussed here. The challenges are not trivial, and these difficulties have presented important barriers to progress in research on the determinants of dementia incidence and progression. The most powerful statistical solutions may require capacity building, including new software and skills development, before they will be broadly adopted. Alongside major investments in strengthening measurements via genotyping, neuroimaging, and biomarker assessments, methodological advances hold promise to accelerate progress toward successful prevention and treatment of dementia. Indeed, without strong research methods, the investments in high-quality measures will be of little use.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Role of the funding source: The MELODEM initiative has been supported by the “Fondation Plan Alzheimer.” J.W. was supported by NIEHS grant R21ES020404 and Alzheimer’s Association grant NIRG-12-242395. M.C.P. was supported by NIA grant T32 AG027668. A.L.G. was supported by NIA grant R03AG045494. S.M.H. was supported by NIA grants P01AG043362 and R01AG026453. M.M.G. was supported by NIA grant R21 AG034385.

References

1. Morris MC, Evans DA, Hebert LE, Bienias JL. Methodological issues in the study of cognitive decline. *Am J Epidemiol.* 1999; 149:789–93. [PubMed: 10221314]
2. Hernan MA, Alonso A, Logroscino G. Cigarette smoking and dementia: Potential selection bias in the elderly. *Epidemiology.* 2008; 19:448–50. [PubMed: 18414087]
3. Glymour MM, Weuve J, Chen JT. Methodological challenges in causal research on racial and ethnic patterns of cognitive trajectories: Measurement, selection, and bias. *Neuropsychol Rev.* 2008; 18:194–213. [PubMed: 18819008]
4. Clouston SA, Kuh D, Herd P, Elliott J, Richards M, Hofer SM. Benefits of educational attainment on adult fluid cognition: International evidence from three birth cohorts. *Int J Epidemiol.* 2012; 41:1729–36. [PubMed: 23108707]
5. Hofer SM, Piccinin AM. Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychol Methods.* 2009; 14:150–64. [PubMed: 19485626]
6. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* 2010; 152:726–32. [PubMed: 20335313]
7. Vandembroucke JP, Elm Ev, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *Ann Intern Med.* 2007; 147:W-163. [PubMed: 17938389]
8. Little J, Higgins J, Ioannidis J, Moher D, Gagnon F, Von Elm E, et al. STrengthening the REporting of Genetic Association studies (STREGA)—An extension of the STROBE statement. *Eur J Clin Invest.* 2009; 39:247–66. [PubMed: 19297801]
9. Kaufman JS, Hernan MA. Epidemiologic methods are useless: They can only give you answers. *Epidemiology.* 2012; 23:785–6. [PubMed: 23038106]
10. Glymour MM. Invited commentary: When bad genes look good -APOE*E4, cognitive decline, and diagnostic thresholds. *Am J Epidemiol.* 2007; 165:1239–46. author reply 47. [PubMed: 17431011]

11. Chodosh J, Seeman TE, Keeler E, Sewall A, Hirsch SH, Guralnik JM, et al. Cognitive decline in high-functioning older persons is associated with an increased risk of hospitalization. *J Am Geriatr Soc.* 2004; 52:1456–62. [PubMed: 15341546]
12. Welmerink DB, Longstreth WT Jr, Lyles MF, Fitzpatrick AL. Cognition and the risk of hospitalization for serious falls in the elderly: Results from the Cardiovascular Health Study. *J Gerontol A Biol Sci Med Sci.* 2010; 65:1242–9. [PubMed: 20584769]
13. Greiner PA, Snowdon DA, Schmitt FA. The loss of independence in activities of daily living: The role of low normal cognitive function in elderly nuns. *Am J Public Health.* 1996; 86:62–6. [PubMed: 8561244]
14. Raji MA, Al Snih S, Ray LA, Patel KV, Markides KS. Cognitive status and incident disability in older Mexican Americans: Findings from the Hispanic established population for the epidemiological study of the elderly. *Ethn Dis.* 2004; 14:26–31. [PubMed: 15002920]
15. Dodge HH, Du Y, Saxton JA, Ganguli M. Cognitive domains and trajectories of functional independence in nondemented elderly persons. *J Gerontol A Biol Sci Med Sci.* 2006; 61:1330–7. [PubMed: 17234830]
16. Yaffe K, Lindquist K, Vittinghoff E, Barnes D, Simonsick EM, Newman A, et al. The effect of maintaining cognition on risk of disability and death. *J Am Geriatr Soc.* 2010; 58:889–94. [PubMed: 20406308]
17. Bassuk SS, Wypij D, Berkman LF. Cognitive impairment and mortality in the community-dwelling elderly. *Am J Epidemiol.* 2000; 151:676–88. [PubMed: 10752795]
18. Euser SM, Schram MT, Hofman A, Westendorp RG, Breteler MM. Measuring cognitive function with age: The influence of selection by health and survival. *Epidemiology.* 2008; 19:440–7. [PubMed: 18414086]
19. Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *J Clin Epidemiol.* 2005; 58:13–9. [PubMed: 15649666]
20. Matthews FE, Chatfield M, Brayne C. Medical Research Council Cognitive F, Ageing S. An investigation of whether factors associated with short-term attrition change or persist over ten years: Data from the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *BMC Public Health.* 2006; 6:185. [PubMed: 16848886]
21. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology.* 2004; 15:615–25. [PubMed: 15308962]
22. Tchetgen Tchetgen E, Glymour M, Weuve J, Shpitser I. To weight or not to weight? On the relation between inverse-probability weighting and principal stratification for truncation by death. *Epidemiology.* 2012; 23:644–6. [PubMed: 22659551]
23. Joly P, Commenges D, Helmer C, Letenneur L. A penalized likelihood approach for an illness-death model with interval-censored data: Application to age-specific incidence of dementia. *Biostatistics.* 2002; 3:433–43. [PubMed: 12933608]
24. Chang CC, Yang HC, Tang G, Ganguli M. Minimizing attrition bias: A longitudinal study of depressive symptoms in an elderly cohort. *Int Psychogeriatr.* 2009; 21:869–78. [PubMed: 19288971]
25. Leffondre K, Touraine C, Helmer C, Joly P. Interval-censored time-to-event and competing risk with death: Is the illness-death model more accurate than the Cox model? *Int J Epidemiol.* 2013; 42:1177–86. [PubMed: 23900486]
26. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, et al. Accounting for bias due to selective attrition: The example of smoking and cognitive decline. *Epidemiology.* 2012; 23:119–28. [PubMed: 21989136]
27. Dufouil C, Brayne C, Clayton D. Analysis of longitudinal studies with death and drop-out: A case study. *Stat Med.* 2004; 23:2215–26. [PubMed: 15236426]
28. Kurland BF, Johnson LL, Egleston BL, Diehr PH. Longitudinal data with follow-up truncated by death: Match the analysis method to research aims. *Stat Sci.* 2009; 24:211. [PubMed: 20119502]
29. Chaix B, Evans D, Merlo J, Suzuki E. Commentary: Weighing up the dead and missing: Reflections on inverse-probability weighting and principal stratification to address truncation by death. *Epidemiology.* 2012; 23:129–31. discussion 132–7. [PubMed: 22157307]

30. Kennelly SP, Lawlor BA, Kenny RA. Blood pressure and the risk for dementia: A double edged sword. *Ageing Res Rev.* 2009; 8:61–70. [PubMed: 19063999]
31. Letenneur L, Gilleron V, Commenges D, Helmer C, Orgogozo JM, Dartigues JF. Are sex and educational level independent predictors of dementia and Alzheimer's disease? Incidence data from the PAQUID project. *J Neurol Neurosurg Psychiatry.* 1999; 66:177–83. [PubMed: 10071096]
32. Luchsinger JA, Patel B, Tang MX, Schupf N, Mayeux R. Measures of adiposity and dementia risk in elderly persons. *Arch Neurol.* 2007; 64:392–8. [PubMed: 17353383]
33. Power MC, Weuve J, Gagne JJ, McQueen MB, Viswanathan A, Blacker D. The association between blood pressure and incident Alzheimer disease: A systematic review and meta-analysis. *Epidemiology.* 2011; 22:646–59. [PubMed: 21705906]
34. McLaughlin T, Buxton M, Mittendorf T, Redekop W, Mucha L, Darba J, et al. Assessment of potential measures in models of progression in Alzheimer disease. *Neurology.* 2010; 75:1256–62. [PubMed: 20921511]
35. Manly JJ, Jacobs DM, Sano M, Bell K, Merchant CA, Small SA, et al. Effect of literacy on neuropsychological test performance in nondemented, education-matched elders. *J Int Neuropsychol Soc.* 1999; 5:191–202. [PubMed: 10217919]
36. Jones RN, Gallo JJ. Education and sex differences in the mini-mental state examination: Effects of differential item functioning. *J Gerontol B Psychol Sci Soc Sci.* 2002; 57:P548–58. [PubMed: 12426438]
37. Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, et al. Reporting standards for studies of diagnostic test accuracy in dementia: The STARDdem Initiative. *Neurology.* 2014; 83:364–73. [PubMed: 24944261]
38. Chmielewski M, Watson D. What is being assessed and why it matters: The impact of transient error on trait research. *J Pers Soc Psychol.* 2009; 97:186–202. [PubMed: 19586248]
39. Lubke GH, Dolan CV, Kelderman H, Mellenbergh GJ. Weak measurement invariance with respect to unmeasured variables: an implication of strict factorial invariance. *Br J Math Stat Psychol.* 2003; 56:231–48. [PubMed: 14633334]
40. Salthouse, TA. *Major issues in cognitive aging.* Oxford: Oxford University Press; 2010.
41. Calamia M, Markon K, Tranel D. Scoring higher the second time around: Meta-analyses of practice effects in neuropsychological assessment. *Clin Neuropsychol.* 2012; 26:543–70. [PubMed: 22540222]
42. Wesnes K, Pincock C. Practice effects on cognitive tasks: A major problem? *Lancet Neurol.* 2002; 1:473. [PubMed: 12849328]
43. Hoffman L, Hofer SM, Sliwinski MJ. On the confounds among retest gains and age-cohort differences in the estimation of within-person change in longitudinal studies: A simulation study. *Psychol Aging.* 2011; 26:778–91. [PubMed: 21639642]
44. Thorvaldsson V, Hofer SM, Johansson B. Aging and late-life terminal decline in perceptual speed: A comparison of alternative modeling approaches. *Eur Psychol.* 2006; 11:196–203.
45. Proust-Lima C, Dartigues JF, Jacqmin-Gadda H. Misuse of the linear mixed model when evaluating risk factors of cognitive decline. *Am J Epidemiol.* 2011; 174:1077–88. [PubMed: 21965187]
46. Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: Choice of the time-scale. *Am J Epidemiol.* 1997; 145:72–80. [PubMed: 8982025]
47. Lamarca R, Alonso J, Gomez G, Munoz A. Left-truncated data with age as time scale: An alternative for survival analysis in the elderly population. *J Gerontol A Biol Sci Med Sci.* 1998; 53:M337–43. [PubMed: 9754138]
48. Thiebaut AC, Benichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: A simulation study. *Stat Med.* 2004; 23:3803–20. [PubMed: 15580597]
49. Amieva H, Le Goff M, Millet X, Orgogozo JM, Peres K, Barberger-Gateau P, et al. Prodromal Alzheimer's disease: Successive emergence of the clinical symptoms. *Ann Neurol.* 2008; 64:492–8. [PubMed: 19067364]

50. Hoffman, L. Considering alternative metrics of time. In: Harring, JR.; Hancock, GR., editors. *Advances in Longitudinal Methods in the Social and Behavioral Sciences* Scottsdale. Scottsdale, Arizona: IAP Information Age Publishing; 2012. p. 255-87.
51. Sliwinski M, Hoffman L, Hofer SM. Evaluating convergence of within-person change and between-person age differences in age-heterogeneous longitudinal studies. *Res Hum Dev.* 2010; 7:45–60. [PubMed: 20671986]
52. Singh-Manoux A, Marmot MG, Glymour M, Sabia S, Kivimaki M, Dugravot A. Does cognitive reserve shape cognitive decline? *Ann Neurol.* 2011; 70:296–304. [PubMed: 21563209]
53. Wilson RS, Beckett LA, Barnes LL, Schneider JA, Bach J, Evans DA, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging.* 2002; 17:179–93. [PubMed: 12061405]
54. Proust C, Jacqmin-Gadda H, Taylor JM, Ganiayre J, Commenges D. A nonlinear model with latent process for cognitive evolution using multivariate longitudinal data. *Biometrics.* 2006; 62:1014–24. [PubMed: 17156275]
55. Hall CB, Lipton RB, Sliwinski M, Stewart WF. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Stat Med.* 2000; 19:1555–66. [PubMed: 10844718]
56. Jacqmin-Gadda H, Commenges D, Dartigues JF. Random change point model for joint modeling of cognitive decline and dementia. *Biometrics.* 2006; 62:254–60. [PubMed: 16542253]
57. Gottesman RF, Schneider AL, Albert M, Alonso A, Bandeen-Roche K, Coker L, et al. Midlife hypertension and 20-year cognitive change: The atherosclerosis risk in communities neurocognitive study. *JAMA Neurol.* 2014; 71:1218–27. [PubMed: 25090106]
58. Power MC, Tchetgen EJ, Sparrow D, Schwartz J, Weisskopf MG. Blood pressure and cognition: Factors that may account for their inconsistent association. *Epidemiology.* 2013; 24:886–93. [PubMed: 24030502]
59. DeBette S, Seshadri S, Beiser A, Au R, Himali JJ, Palumbo C, et al. Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. *Neurology.* 2011; 77:461–8. [PubMed: 21810696]
60. Whitmer RA, Quesenberry CP, Zhou J, Yaffe K. Timing of hormone therapy and dementia: The critical window theory revisited. *Ann Neurol.* 2011; 69:163–9. [PubMed: 21280086]
61. O'Brien J, Jackson JW, Grodstein F, Blacker D, Weuve J. Postmenopausal hormone therapy is not associated with risk of all-cause dementia and Alzheimer's disease. *Epidemiol Rev.* 2014; 36:83–103. [PubMed: 24042430]
62. Ritchie CW, Ritchie K. The PREVENT study: A prospective cohort study to identify mid-life biomarkers of late-onset Alzheimer's disease. *BMJ Open.* 2012; 2
63. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med.* 1975; 4:518–25. [PubMed: 1208363]
64. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med.* 2013; 32:1584–618. [PubMed: 23208861]
65. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000; 11:550–60. [PubMed: 10955408]
66. Khachaturian AS, Meranus DH, Kukull WA, Khachaturian ZS. Big data, aging, and dementia: Pathways for international harmonization on data sharing. *Alzheimers Dement.* 2013; 9:S61–2. [PubMed: 24125464]
67. Montine TJ, Koroshetz WJ, Babcock D, Dickson DW, Galpern WR, Glymour MM, et al. Recommendations of the Alzheimer's disease-related dementias conference. *Neurology.* 2014; 83:851–60. [PubMed: 25080517]
68. Qizilbash N, Gregson J, Johnson ME, Pearce N, Douglas I, Wing K, et al. BMI and risk of dementia in two million people over two decades: A retrospective cohort study. *Lancet Diabetes Endocrinol.* 2015; 3:431–6. [PubMed: 25866264]
69. Gustafson D. BMI and dementia: Feast or famine for the brain? *Lancet Diabetes Endocrinol.* 2015; 3:397–8. [PubMed: 25866263]
70. Piccinin, A.; Hofer, S. Integrative analysis of longitudinal studies of aging: collaborative research networks, meta-analysis, and optimizing future studies. In: Hofer, S.; Alwin, D., editors.

Handbook of cognitive aging: Interdisciplinary perspectives. Thousand Oaks, California: Sage Publications; 2008. p. 446-76.

71. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA*. 1996; 276:637–9. [PubMed: 8773637]
72. Grammer TB, Hoffmann MM, Scharnagl H, Kleber ME, Silbernagel G, Pilz S, et al. Smoking, apolipoprotein E genotypes, and mortality (the Ludwigshafen Risk and Cardiovascular Health study). *Eur Heart J*. 2013; 34:1298–305. [PubMed: 23382465]

RESEARCH IN CONTEXT

1. Systematic review: Clinical and population research on dementia and related neurologic conditions, including Alzheimer’s disease, faces several unique methodological challenges. Progress to identify preventive and therapeutic strategies rests on valid and rigorous analytic approaches, but the research literature reflects little consensus on “best practices.” There is a great need for guidelines addressing specific issues to dementia/cognitive ageing research, in line to those developed through the Enhancing the QUALity and Transparency Of health Research (Equator) network.
2. Interpretation: Several methodological challenges arise in studies of the determinants of dementia risk and cognitive decline. Currently, researchers handle these challenges differently, making it difficult to directly compare studies and combine evidence. Guidelines were generated through the international initiative MELODEM (MEthods for LOngitudinal studies in DEMentia) to outline a set of methodological problems that should routinely be addressed in dementia research.
3. Future directions: We hope researchers adopting the MELODEM guidelines will routinely acknowledge these challenges and justify analytic decisions. We anticipate that MELODEM will provide a platform for continued discussion and innovation of methodological tools to strengthen dementia research.

- **Clarify the question:** State the scientific question of interest, specifying whether it is causal or predictive; identify the relevant population; define the estimand targeted by the statistical analysis; and explain the assumptions under which the analytic approach answers this question. (Section 2.1)
- **Dropout:** Describe dropout and correlations of dropout with exposure and past outcome (e.g., cognitive) history. Specify how potential dropout bias was handled. Use sensitivity analyses to assess the extent to which results could be influenced by informative dropout. (Section 2.2; Supplementary Appendix 1, § 1)
- **Survival:** Describe survival and correlations of survival with exposure and past cognitive history. Specify how potential survival bias was handled. If exposure occurred prior to study baseline (e.g., education or childhood exposures), estimate the fraction of exposed and unexposed individuals who survived to study enrollment. (Section 2.2; Supplementary Appendix 1, § 2)
- **Selective study participation:** Describe fraction of invited individuals who consented to study participation, how participation rates differed by known background characteristics, and how selective participation was handled in analyses. If exposure occurred prior to study baseline and data are available, estimate the fraction of exposed and unexposed individuals who agreed to enroll. (Section 2.2; Supplementary Appendix 1, § 3)
- **Practice effects:** Describe any evidence of practice effects on the outcome measure and how potential practice effects were handled in the analysis. In longitudinal models, assess whether primary analysis results differ when including an indicator for first assessment. (Section 2.3, § 1; Supplementary Appendix 2, § 3)
- **Reliability:** Describe the reliability of the primary exposure and outcome measures. (Section 2.3, § 4; Supplementary Appendix 2, § 2)
- **Unequal interval scaling:** If the primary outcome is a continuous measure without a meaningful absolute unit scale (such as cognition, or depression), show its distribution and describe any evidence that problems with interval scaling may be present. If possible, describe likely direction of bias due to unequal interval scaling, including ceiling and floor effects. (Supplementary Appendix 2, § 4)
- **Time scale:** Specify and justify the time scale used. Whatever the time scale, adjust the analysis for age at a given time (e.g., age at entry). When the appropriate time scale is uncertain, assess sensitivity of primary study results to alternatives. (Section 2.4; Supplementary Appendix 3, § 1 & 2)
- **Nonlinear cognitive trajectory:** Explore the possibility that cognitive decline is nonlinear, and if it is, analytically account for the nonlinearities. (Section 2.4; Supplementary Appendix 3, § 3)
- **Time-varying exposures:** Specify whether exposure was time-updated and, if so, specify if any other operationalization of the exposure was considered besides the most-recent value. Justify choice of timing of exposures and confounders used in analyses. Interpret estimates of association in light of the exposure timing used. (Section 2.5; Supplementary Appendix 4, § 1)
- **Time-varying confounding:** Assess whether variables treated as confounders partially mediate effects of interest and distinguish time-varying from time-constant confounders. (Section 2.5; Supplementary Appendix 4, § 3)
- **High dimensional data:** Account for false discovery. Correct for overfitting. (Section 2.6, § 4)

Fig. 1.
Guidelines for reporting methodological challenges and evaluating bias in cognitive decline and dementia research.

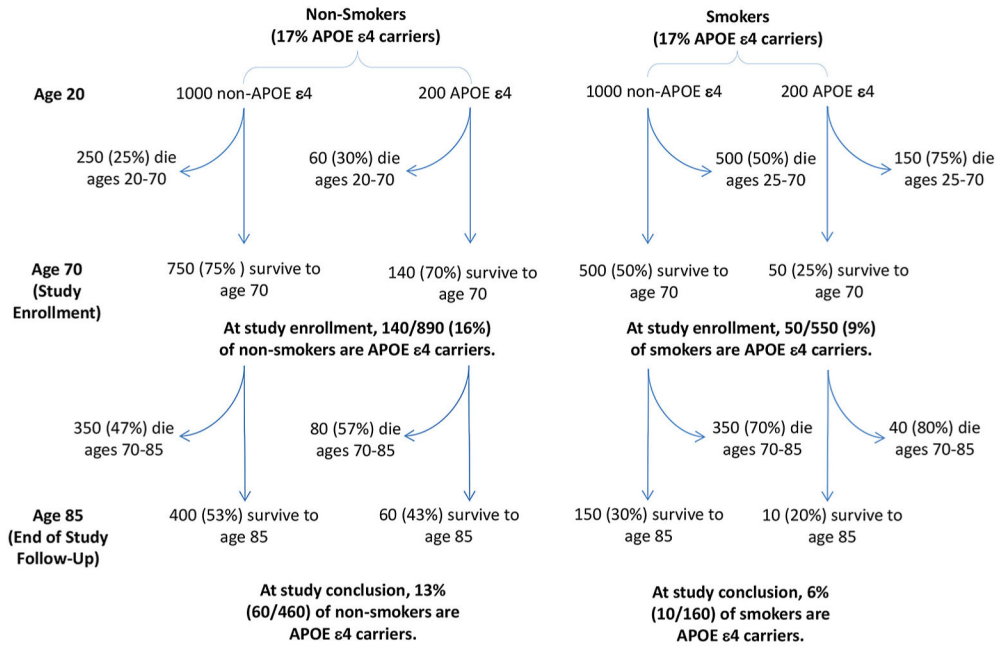


Fig. 2. Hypothetical illustration of selection processes before and after study enrollment. At age 20, smoking and *APOE* ε4 are unrelated, but these risk factors synergistically affect mortality, with more than multiplicative effects on survival up to age 70 [72]. By the time of study initiation at age 70, smokers are very unlikely to be *APOE* ε4 carriers. Analyses that did not control for *APOE* ε4 would conflate *APOE* status and smoking and spuriously underestimate effects of smoking.

Table 1

Selection processes: problems and commonly adopted analytic approaches*

Differential attrition of enrolled participants	Differential survival of enrolled participants	Differential study enrollment or “muting”
<i>Approaches commonly applied to multiple selection problems</i>		
Ignore	Ignore	Ignore
Sensitivity analyses for magnitude of bias under plausible set of selection processes	Sensitivity analyses for magnitude of bias under plausible set of selection processes	Sensitivity analyses for magnitude of bias under plausible set of selection processes
Assess bounds based on best or worst case assumptions	Assess bounds based on best or worst case assumptions	Assess bounds based on best or worst case assumptions
Model determinants of selection to evaluate whether ignoring selection is appropriate	Model determinants of selection to evaluate whether ignoring selection is appropriate	Model determinants of selection to evaluate whether ignoring selection is appropriate
Adjust for determinants of selection	Adjust for determinants of selection	Adjust for determinants of selection
Weight on the inverse of the probability of selection	Weight on the inverse of the probability of selection	Weight on the inverse of the probability of selection
Instrumental variable methods, if an instrument for selection is available	Instrumental variable methods, if an instrument for selection is available	Instrumental variable methods, if an instrument for selection is available
Joint modeling of selection process (dropout, death, enrollment) and outcome	Joint modeling of selection process (dropout, death, enrollment) and outcome	Joint modeling of selection process (dropout, death, enrollment) and outcome
<i>Approaches commonly applied only to specific selection problems</i>		
Multiple imputation and likelihood-based estimation including covariates related to missingness mechanism	Competing risks analysis (only when dementia is the outcome)	Principal stratification

* For many of these approaches, there is currently limited empirical or theoretical evidence comparing the performance (i.e., providing a precise estimate of the effect of interest) in dementia research.

Table 2

Measurement challenges: problems and commonly adopted analytic approaches*

Validity of measurement	Reliability/random measurement error	Practice or retest effects	Unequal-interval scaling (including ceilings/floors on measures)
<i>Approaches commonly applied to multiple measurement problems</i>			
Ignore	Ignore	Ignore	Ignore
Multivariate latent variable methods or measurement error models	Multivariate latent variable methods or measurement error models		
<i>Approaches applied only to specific measurement problems</i>			
Compare to a gold standard/criterion validity	Instrumental variable analyses	Drop the first assessment or average first two assessments	Drop observations at the ceiling/ floor or otherwise condition on the baseline score
Compare to measures of theoretically correlated variables	Use composite scores from multiple neuropsychological assessments (e.g., summed Z-scores [†])	Choose tests with limited retest effects	Item response theory or factor analysis based models. Factor analyses, imposing distributional assumptions
Evaluate Differential Item Functioning (DIF) and implement statistical corrections or adjustment for source of DIF		Randomize time of first assessment	Rescale by Z-scoring [†]
		Indicator for first assessment	Transform the measure with a monotonic transformation intended to reduce non-interval scales (e.g. logarithm, box-cox, specifically designed normalizing transformation)
		Other models of practice (linear or non-linear increases in practice effects)	Categorize the outcome (impaired vs. not impaired)
		Mixed models identifying practice effects based on time-varying interview delays	Tobit regression models (for ceilings/floors) or quantile regressions
			Joint estimation of a normalizing transformation of the outcome and the coefficients

* For many of these approaches, there is currently limited empirical or theoretical evidence comparing the performance (i.e., providing a precise estimate of the effect of interest) in dementia research.

[†] Z-scoring rescales each individual's raw score with respect to the distribution of scores for other individuals in the sample. From each individual's raw score, the Z-score is calculated by subtracting the sample mean (usually at baseline) and dividing by the sample standard deviation (also at baseline).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Defining the time scale for longitudinal analyses: problems and commonly adopted analytic approaches*

Divergence of within-person change and between-person age differences	Analysis of terminal decline preceding death, dementia, or other “milestone” events	Nonlinear cognitive trajectories
<i>Approaches commonly applied to multiple time scale problems</i>		
Ignore	Ignore	Ignore
<i>Approaches applied only to specific time scale problems</i>		
Age as the time-scale with adjustment for age at entry or time-from-entry as the time-scale adjusting for age at entry	Analysis among the participants who had the event	Polynomial trajectory (quadratic, cubic)
Use of age at assessment as the time scale, without adjustment for age at entry.	Time to event as time scale in the group with event versus time to last measure for the healthy participants matched by or adjusted for the age at the last measure among others	Trajectories with random, pre-specified, or empirically selected change-points
Other time scale of interest adjusting for a cross-sectional age (possibly other than age at entry)	Joint model of the longitudinal outcome and the time to the event of interest (death, dementia or others)	Flexible parametric (splines, fractional polynomials) or non-parametric trajectories

* For many of these approaches, there is currently limited empirical or theoretical evidence comparing the performance (i.e., providing a precise estimate of the effect of interest) in dementia research.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Handling time-varying exposures and time-varying confounding: problems and commonly adopted analytic approaches *

Time-varying exposures	Time-varying confounding
<i>Approaches commonly applied to both time-varying exposures and time-varying confounding</i>	
Ignore (e.g., consider variable at a single point in time)	Ignore (e.g., consider variable at a single point in time)
Marginal structural models and inverse probability weighting Structural nested models	Marginal structural models and inverse probability weighting Structural nested models
<i>Approaches commonly applied to either time-varying exposures or time-varying confounding</i>	
Time-to-event models, allowing exposure to update or lag	Compare effect estimates with or without adjustment for time-varying confounders
Summaries of time-varying exposure (e.g. average, duration, age at initiation)	Longitudinal propensity score models
Compare estimates from several models using exposure status at a single point in time or moving time windows; formally test alternative lifecourse models.	Instrumental variables models

* For many of these approaches, there is currently limited empirical or theoretical evidence comparing the performance (i.e., providing a precise estimate of the effect of interest) in dementia research.

Table 5

Handling high dimensional data: challenges and commonly adopted analytic approaches

Multiple comparisons/false discovery	Summarizing multiple highly correlated variables	Regression with high dimensional data
Family wise error correction (e.g. Bonferroni)	Theoretically motivated summaries or selected indicators based on prior knowledge, e.g., candidate gene approaches	Preselection of the variables of interest for adjustment
False discovery rate (e.g. BH correction)	Combination of variables (e.g. principal components analysis, partial least square)	Regularization methods (e.g. Lasso, ridge regression, elastic net)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript