

---

## Review

---

# Closing the Gap Between Observational Research and Randomized Controlled Trials for Prevention of Alzheimer Disease and Dementia

Melinda C. Power\*, Brittany C. Engelman, Jingkai Wei, and M. Maria Glymour

\* Correspondence to Melinda C. Power, Milken Institute School of Public Health, George Washington University, 950 New Hampshire Avenue, NW, Washington, DC 20052 (e-mail: power@gwu.edu).

*Accepted for publication April 4, 2022.*

---

Although observational studies have identified modifiable risk factors for Alzheimer disease and related dementias (ADRD), randomized controlled trials (RCTs) of risk factor modification for ADRD prevention have been inconsistent or inconclusive. This finding suggests a need to improve translation between observational studies and RCTs. However, many common features of observational studies reduce their relevance to designing related RCTs. Observational studies routinely differ from RCTs with respect to eligibility criteria, study population, length of follow-up, treatment conditions, outcomes, and effect estimates. Using the motivating example of blood pressure reduction for ADRD prevention, we illustrate the need for a tighter connection between observational studies and RCTs, discuss barriers to using typically reported observational evidence in developing RCTs, and highlight methods that may be used to make observational research more relevant to clinical trial design. We conclude that the questions asked and answered by observational research can be made more relevant to clinical trial design and that better use of observational data may increase the likelihood of successful, or at least definitive, trials. Although we focus on improving translation of observational studies on risk factors for ADRD to RCTs in ADRD prevention, the overarching themes are broadly applicable to many areas of biomedical research.

antihypertensive agents; blood pressure; cognition; dementia; epidemiologic methods; observational studies; randomized controlled trials

---

Abbreviations: ADRD, Alzheimer disease and related dementia; ARIC, Atherosclerosis Risk in Communities Study; ATE, average treatment effect; BP, blood pressure; HRS, Health and Retirement Study; ITT, intention-to-treat; RCT, randomized controlled trial; SBP, systolic blood pressure.

---

## INTRODUCTION

Globally, the number of older adults with Alzheimer disease or related dementias (ADRD) will reach approximately 115 million by 2050 (1, 2). The personal and societal impacts of dementia are enormous (3–5), and the global cost associated with dementia was estimated to be US\$818 billion in 2015 (6). Currently, we have no treatments to stop or slow the progression of ADRD (1, 2, 7).

From 2002 to 2020, only 1 new drug was approved by the US Food and Drug Administration for the treatment of Alzheimer disease, reflecting a failure rate of randomized controlled trials (RCTs) of pharmaceutical interventions of 99.6% (8). Thus, efforts to address the ADRD epidemic increasingly focus on risk-factor modification (9–13). However, evidence across ADRD risk-factor modification

prevention trials has been inconsistent or inconclusive, precluding strong conclusions about the efficacy of any particular strategy (14).

RCTs of risk-factor modification are typically fielded on the basis of observational evidence. Lack of conclusive trials suggests a need for improving translation of observational studies on risk factors for ADRD to RCTs of risk-factor modification for ADRD prevention. Currently, many common features of observational studies reduce their relevance to designing RCTs. Observational studies routinely differ from RCTs with respect to eligibility criteria, study population, length of follow-up, treatment conditions, outcomes, and estimated effect. As a result, questions asked and answered by observational studies often have little relevance to the decisions that need to be made when designing an RCT.

In this narrative review, we use the motivating example of blood pressure (BP) and ADRD to illustrate the need for a tighter connection between observational studies and RCTs, discuss barriers to using typically reported observational evidence in developing RCTs, and highlight methods that may be used to make observational research more relevant to clinical trial design. We conclude that the questions asked and answered by observational research can be made more relevant to clinical trial design and that better use of observational data can increase the likelihood of definitive, if not successful, trials. Although we focus on improving translation of observational studies on risk factors for ADRD to RCTs in ADRD prevention, the overarching themes are broadly applicable to many areas of biomedical research.

### BP MANAGEMENT FOR ADRD PREVENTION

Collectively, the existing literature suggests an age-dependent association between BP and late-life cognitive health. Midlife hypertension appears to confer excess risk of later-life cognitive decline, cognitive impairment, and dementia, whereas there is little evidence that hypertension in later life leads to cognitive loss (15–17). Although the story is likely somewhat more complicated—for example, BP may differentially affect multiple dementia-related pathogenic processes (18), differences in duration of hypertension, rather than age at onset, may account for the age-dependent pattern of findings (19), and specific patterns of BP over time may be particularly detrimental (20, 21)—the weight of the observational evidence supports the idea that better BP management, particularly in midlife, will lower risk of cognitive decline and dementia.

Over the past 30 years, multiple RCTs have evaluated the effect of antihypertensive medication use on cognitive health (Web Table 1, available at <https://doi.org/10.1093/epirev/mxac002>). Taken individually, these trials have not demonstrated that intervening to lower BP through use of antihypertensive medications prevents cognitive impairment or dementia during follow-up. However, many RCTs were halted early, and authors report protective point estimates with confidence intervals that include the null or report positive results only for secondary cognitive outcomes. Thus, although the trial results do not demonstrate that antihypertensive medication use is an effective strategy for ADRD prevention, they also do not definitively demonstrate that it is an ineffective strategy. For example, in the Systolic Blood Pressure Intervention Trial, Memory and Cognition in Decreased Hypertension (SPRINT MIND) study, researchers evaluated whether assignment to intensive lowering of systolic blood pressure (SBP) to a target of 120 mmHg versus standard lowering of SBP to a target of 140 mmHg affected dementia risk (22, 23), but the trial was stopped early for demonstrated benefit on cardiovascular health. Ultimately, SPRINT MIND did not show a benefit of the intervention on dementia risk (hazard ratio = 0.93, 95% confidence interval: 0.73, 1.18), but researchers reported a protective effect on a secondary outcome, mild cognitive impairment (MCI; hazard ratio = 0.83, 95% confidence interval: 0.70, 0.99) (22, 23).

Given lack of clear evidence from individual trials, meta-analyses of findings from multiple RCTs are appealing. However, conclusions drawn from meta-analytic efforts vary; authors disagree on whether different trials can be meta-analyzed and, if so, which trials meet criteria for inclusion in the meta-analysis, including criteria related to risk of bias (24–27). Therefore, definitive conclusions elude the field.

Fielding RCTs that do not provide definitive answers can cause harm. RCTs are extremely expensive and take years to conduct. For example, the amounts of the awarded federal contracts funding the Systolic Blood Pressure Intervention Trial, the parent of SPRINT MIND, exceed \$120 million (28, 29). Public perception and trust in biomedical research are also at stake, because we have spent time, money, and energy on an endeavor that did not yield definitive conclusions. Trial participation is also at risk. Why would potential participants volunteer to participate in an RCT if they do not think the knowledge to be gained is worth the risks and inconvenience of participation?

Although the best available observational evidence showed a convincing link between lower BP and reduced dementia risk, RCTs of BP lowering for ADRD prevention fielded on this evidence have been neither successful nor conclusive. These findings suggest a need to revise our approach to future ADRD prevention trials to maximize likelihood of successful, or at least definitive, trials.

### DIFFERENCES BETWEEN OBSERVATIONAL STUDIES AND RCTS

Better use of observational studies to inform choices about clinical trial design may increase the likelihood of successful trials and definitive results. However, most observational studies do not currently provide information needed to effectively guide clinical trial design. Using the example of BP management and ADRD prevention, we consider differences in the composition of the study population, length of follow-up, treatment and exposure conditions, outcomes, and analytic approaches across observational studies and related RCTs, and we discuss how observational researchers can alter their approach to be more relevant to RCT design (Table 1).

#### Eligibility criteria and sample composition

Although not universal, observational studies typically have few restrictions on participation beyond age or locality requirements and often use probability sampling methods to recruit samples representative of the target population. For example, the Atherosclerosis Risk in Communities (ARIC) Study and the Rotterdam Study have both been used to investigate the link between BP and late-life cognitive or brain health (20, 30–36). In ARIC, random probability sampling from available registers was used to recruit participants representative of 4 US communities; eligibility criteria required only that participants be noninstitutionalized, community-dwelling adults aged 45 to 64 years (37). Similarly, the Rotterdam study recruited residents of 1 district of Rotterdam using random sampling from an existing register, with eligibility criteria requiring a minimum age of 55 years (38).

**Table 1.** Comparison of Randomized Controlled Trials and Observational Study Design Features

Feature	Common in RCTs	Common in Observational Studies	How to Make Observational Studies More Relevant to RCT Design
Eligibility criteria	Narrow, with a defining risk factor	Few restrictions on study participants	Compare estimated effects within subgroups, particularly those aligning with particular risk strata or potential for disproportionate benefit
Sample composition	Convenience samples, with potential for highly selected samples	Clinical or population-based samples	Systematically evaluate effect modifiers and report subgroup effect estimates; reweight estimates or apply restriction to explore effects in samples with different composition
Length of follow-up	Relatively brief, with intervention often lasting for only ~12–48 months; stopped when continuation would result in harm to 1 group	Feasible and common to follow participants for years to decades	Evaluate how exposures relate to subsequent outcomes in the short, medium, or long term; quantify the likelihood and timing of other benefits and harms of change in exposure that might lead to termination of treatment in a trial (e.g., evidence of benefit or harm on a different outcome)
Treatment conditions	Test intervention vs. control condition (e.g., a specific medication regimen); intervention condition is chosen based on expectation it will lead to a desired risk factor status	Contrast groups with given levels or intensity of an exposure (e.g., presence or absence of a risk factor at a given time)	Evaluate the impact of treatments that affect risk-factor status or changes in exposure using methods that address issues of confounding by indication and reverse causation; evaluate heterogeneity in estimated effects of changes in exposure based on the presumed cause of exposure change (e.g., medication, lifestyle change, prevalent disease); use study designs that exploit sources of variation in exposure that are unlikely to be confounded (e.g., quasi-experimental study designs)
Outcome	Primary outcome prespecified, sometimes with governmental or marketing approval in mind	Based on outcome measure available in the data and most statistically significant	Systematically report results for all available measures, regardless of statistical significance, to identify heterogeneity in strength of effect
Analysis	Analyses are prespecified; intent-to-treat analysis estimating the average marginal treatment effect of randomization	Analytic options are endless; conventional analysis is often a conditional effect estimate comparing those with or without a given exposure	Report average marginal effect of exposure (contrasting everyone vs. nobody exposed) to provide bounds on potential effect sizes, recognize or estimate attenuation due to nonadherence and bias due to misclassification

Abbreviation: RCT, randomized controlled trial.

As such, few people are systematically excluded from participation, and the resulting sample should be representative of and generalizable to the population sampled (although it is important to acknowledge that participant refusal or selective attrition may lead to a sample that is not truly representative (39–41)).

RCTs are typically convenience samples because they rely on accessible patient populations and advertising to recruit participants. As such, they are not necessarily representative of or generalizable to the defined target population of interest. They also typically have much narrower eligibility criteria. In addition to age and study setting, inclusion criteria are designed to exclude persons without indications for treatment, exclude those for whom the effect of treatment is known, and to enrich the sample with those at high

risk of the outcome (Web Table 1). For example, SPRINT MIND required participants to be at least 50 years old, and to have no history of diabetes or stroke, high risk of cardiovascular events, and SBP of 130–180 mmHg, with more specific SBP criteria based on the number of antihypertensive medications taken (22, 23). RCTs also frequently impose multiple exclusion criteria, based on logistical or other concerns. For example, exclusion criteria for SPRINT MIND reflect safety or feasibility concerns and select for persons likely to adhere to the protocol and complete follow-up (e.g., indication for a specific antihypertensive medication, arm circumference incompatible with accurate BP readings using standard equipment, expected survival of <3 years, active alcohol abuse, lack of support from a primary health provider, history of missed clinic visits).

Observational studies can be used to identify persons who are most likely to benefit from a given intervention through estimation of subgroup effects. This information can then be used to set inclusion and exclusion criteria that focus the resulting trial on those most likely to benefit. For example, it may be useful to stratify the observational sample based on risk profiles or other characteristics. Observational researchers often avoid reporting extensive subgroup analyses, given legitimate concerns about power and the likelihood of false-positive results that result from “data fishing.” Although a single observational study may not be adequately powered to formally test for effect modification, estimates within any single data set may lack desired precision, and subgroup estimation can increase type I error, it is rare that there is only 1 data set in which an analysis could be performed. Routine reporting of subgroup analyses across multiple studies would provide trial planners a source of evidence on relevant heterogeneity in treatment effects and ensure that chance findings are disproven by subsequent research.

Observational studies could also be used to provide an assessment of whether an effect is likely to be observable given the eligibility constraints of RCTs. Although there is growing interest in using the transportability framework to use trial results to estimate effects in external, real-world populations (42–45), application of this concept in reverse may also allow observational researchers to provide insight into the expected outcome of RCTs in samples that more closely resemble those who would meet criteria for inclusion in a trial. Given the large amount of data observational studies typically obtain, it should be possible to apply restriction, weighting, or transportability algorithms to observational data to estimate associations in samples that resemble trial populations. Such analyses would provide insight into whether trial-eligibility criteria influence the likelihood of a positive trial and can be used to explore whether it is worth focusing on a specific target population, given safety and feasibility constraints.

### Length of follow-up

Observational epidemiologic studies of risk factors for disease often use data collected across decades of follow-up; for example, analyses in the North Karelia/FINMONICA cohorts reporting that midlife hypertension increases dementia risk were based on a mean 21 years of follow-up (46). Similar findings in ARIC were based on a median 23 years of follow-up (30). To the contrary, RCTs rarely follow participants for longer than 5 years. For example, mean follow-up for dementia was 5 years in the Systolic Hypertension in the Elderly Program (SHEP) trial (47) and 2.2 years in the Hypertension in the Very Elderly Trial Cognitive Function Assessment (HYVET-COG) trial (48).

The relatively short follow-up period for RCTs is unlikely to change. Unlike observational studies, for which length of follow-up is limited only by continued funding and participant goodwill, the length of RCTs is constrained by safety and feasibility concerns. RCTs also do not always achieve planned follow-up. For example, RCTs are stopped early if there is demonstrated benefit or harm. This can be a barrier

to definitive results for sub-studies or secondary outcomes, which may only be adequately powered under full follow-up. This was an issue in trials of pharmacologic blood-pressure lowering for ADRD prevention. For all of the trials listed in Web Table 1, the primary goal of each RCT was to establish benefit or harm on a cardiovascular outcome; cognition or dementia were secondary outcomes or were primary outcomes for a sub-study. Many of the trials were stopped early because the intervention was shown to have benefit on cardiovascular health. This made it impossible to ethically continue to study the effect of receiving the intervention on cognition across the intended duration of follow-up.

Although extended follow-up of trials is becoming more common (e.g., ACCORDION MIND (49); SPRINT MIND 2.0 (50)), such follow-up essentially turns the trial into an observational study. At the end of the trial, results are made publicly available, treatment assignment is unblinded, and all participants are encouraged to change to the new standard of care identified by the trial; persons followed from this point forward are no longer randomly assigned or blinded to a treatment condition. Thus, although extended follow-up may make it possible to quantify the long-term, intention-to-treat (ITT) effect of a short period of intervention, it does not allow estimation of the ITT effect of long-term, sustained intervention.

Because observational studies often follow participants for years to decades, they may be able to provide useful insight into the expected effect of an exposure at a given age over a variety of follow-up periods. Conceptual models and associated methods from the field of environmental epidemiology may be particularly relevant to this task (e.g., see the study of Wang et al. (51)). Such models aim to identify etiologically relevant periods or patterns of exposure, while recognizing that most exposures are ubiquitous but of varying intensity, that there may be critical windows in the life course when exposure is particularly harmful, and that there is often a lag between exposure and an observable effect on the outcome. This framework can be applied to observational data in an attempt to identify health effects of an exposure over multiple periods (short, medium, long term), and can be used to provide an assessment of whether we would expect to see a benefit of intervention within a feasible follow-up period. Work of this nature has been done in the area of BP and cognition. Observational studies with short follow-up are mostly null, whereas studies of midlife hypertension and late-life cognition—which, by definition, have decades of follow-up—most reliably report an adverse association (15). Combined with evidence that increasing duration of hypertension is relevant to cognition (19), these observational studies suggest that short-term interventions on BP are unlikely to provide measurable reductions in dementia risk. Given this, it is unsurprising that the existing RCTs of BP reduction for ADRD prevention do not demonstrate benefit within 2 to 5 years of follow-up.

Many interventions have multiple plausible effects, and trials often explicitly specify secondary endpoints of interest. If effects on a primary outcome are established (e.g., cardiovascular events), it is no less important to understand the impact of intervention on other outcomes (e.g., dementia).



A full picture is required to identify optimal treatments, because treatment decisions initiated because of benefits on 1 health outcome may be bolstered or balanced by benefits or harms to a second outcome. Observational studies can help us design trials that are less likely to have inconclusive findings for secondary outcomes when trials are terminated early. Most observational studies collect data on a wide variety of participant characteristics, health behaviors, and health outcomes; as such, they can be used to predict the likely and varied benefits and harms of an intervention. Because they are not limited to a primary outcome, observational studies can be used to evaluate the likelihood of early termination due to other benefits or harms—essentially accounting for the competing risks of the other likely outcomes in the trial—allowing for a more informed approach to the design of RCT sub-studies. For example, if findings from observational studies suggest it is reasonably likely that we would obtain evidence of benefit or harm on the primary outcome before the end of the full, planned, follow-up period for a given trial, it may be justifiable to power sub-studies assuming shortened follow-up, to ensure enrollment and follow-up for sub-studies begin at the same time as primary enrollment, and/or to choose a clinically relevant, sensitive, intermediate endpoint (e.g., cognitive change) as the secondary or sub-study outcome of interest.

### Exposure or treatment conditions

The treatments tested in trials often bear little resemblance to the exposures contrasted in the observational studies used to motivate trials. Observational studies investigating risk factors for disease typically contrast persons with and without a given exposure or, alternately, persons with differing intensities of exposure. For example, observational studies of BP on dementia test whether risk of dementia differs across those with and without hypertension or across those with different SBPs at a defined point in time. To the contrary, RCTs of risk-factor modification for disease prevention typically test a treatment that affects the risk factor identified in observational studies. For example, trials of BP lowering for ADRD prevention typically test the impact of receiving a specific antihypertensive medication (vs. a placebo) or of being assigned to pharmacologic treatment with the goal of reaching a target BP (vs. an alternate target) (Web Table 1).

As such, RCTs test the effect of randomization to an intervention that might achieve a given level of a risk factor (e.g., antihypertensive medication use). This is potentially quite different from the impact of having a given level of a risk factor (e.g., SBP of 145 mmHg), or even the impact of a change to a given level of a risk factor (e.g., decline in SBP from 160 mmHg to 145 mmHg) for several reasons. Assignment to an intervention in an RCT may or may not produce the expected risk-factor change. For example, in HYVET-COG, the difference in BP achieved between the intervention and control groups was substantial (SBP 15.0 mmHg, diastolic BP 5.9 mmHg) (48), whereas in the Study on Cognition and Prognosis in the Elderly (SCOPE), the difference was minimal (SBP 3.2 mmHg, diastolic BP 1.6 mmHg) (52). Moreover, treatments tested in RCTs may influence outcomes through

multiple mechanisms. For example, antihypertensive medication use may influence ADRD risk through effects of the medication on brain calcium regulation, modulation of the renin–angiotensin system, or anti-inflammatory effects, in addition to any effects through lowering BP (53). These other effects would not be captured in an observational study that contrasted risk of ADRD across studies showing differences in BP. As such, differences in risk across groups defined by risk-factor status in an observational study may not directly correspond to the effect of a treatment intended to change risk-factor status in a trial.

In addition, exposures are often correlated with each other and across time. This raises questions about whether the association between a given exposure and outcome in an observational study reflects a causal effect of that exposure at that time on the outcome. First, the exposure of interest may not be the causal agent driving the observed association in the observational study. For example, residual confounding may lead to an association that does not reflect a causal effect. Second, risk-factor status is correlated over time within an individual. Therefore, the fact that we observe an association between risk-factor status at a given time and an outcome in an observational study does not necessarily mean that that exposure at that time affects the outcome. For example, an individual who is hypertensive at age 70 years was very likely hypertensive at age 55 years and will likely remain hypertensive at age 90 years. They may also be more likely to experience episodes of extremely high BP. If BP affects risk of ADRD through the cumulative effects of decades of hypertension or through repeated episodes of extremely high BP, we should see an association between hypertension at age 70 years and dementia risk in an observational study. However, in this case, we would not expect that achieving a nonhypertensive target BP level for 1 year at the age of 70 years (e.g., in the context of an RCT) would produce a difference in dementia risk equivalent to the association between hypertension at age 70 years and incident dementia observed in the observational study.

Observational studies can better inform clinical trials by providing exposure contrasts that align with treatments we can test in a trial. In the context of BP and risk of ADRD, careful evaluation of the impact of treatments known to affect BP (e.g., antihypertensive medication use) are warranted. That said, naïve comparisons of groups with and without the treatment are unlikely to provide sufficient information to guide trial design. For example, simple comparisons of ADRD risk across persons with and without antihypertensive medication use in observational data will not recover the causal effect of antihypertensive medication use on dementia risk, because of reverse causation (e.g., cognitive impairment leads to poor adherence or de-prescribing of antihypertensive medications), confounding by indication (e.g., persons with antihypertensive medication use are more likely than their nonmedicated counterparts to have prior hypertension and clinical or subclinical cardiovascular disease, which increase ADRD risk), and interpretational difficulties due to correlation of medication status over time (i.e., medication status at a single time point is unlikely to be the causal agent, just as SBP at age 70

years was not likely to be the causal agent, but is correlated with it). Instead, careful analyses, using methods designed to address these issues (e.g., restriction to persons with an indication for medication use, new user designs, propensity score matching, active comparators) are necessary (54–56). However, it is important to note that careful analyses may not be sufficient. For example, if reverse causation is likely, it may be impossible to estimate short-term effects of a potential intervention in observational data.

Simple comparisons of persons with and without exposure change may also be useful but will suffer from similar issues. In addition, they are likely to be confounded by measured and unmeasured factors and will be difficult to interpret without full understanding of the causes of exposure change (e.g., BP can change because of heart failure, atherosclerosis, exercise, diet, medication use). Analyses that evaluate heterogeneity in the estimated effect of changes in exposure based on the presumed cause of exposure change may be particularly valuable. However, this presents a challenge, because data sets with sufficient sample size (e.g., claims data) are unlikely to have sufficient information to identify causes of exposure change and adequately control for confounding.

It should also be noted that observational study analyses designed to estimate the effect of treatments or exposure change may be limited by absence of an adequate comparison group. If all of those with specific indications for treatment receive treatment, or if everyone in a group defined by participant characteristics experiences similar exposure change, there is no one comparable to contrast them with, and it will not be possible to estimate a causal effect of exposure in this group (57).

In the face of these substantial challenges, observational studies that take advantage of natural experiments that create variation in exposure, where variation is unlikely to be strongly confounded, provide another useful and complementary approach (58). For example, changes to treatment guidelines for SBP in nationalized health care settings may provide an opportunity to compare those just above the threshold needed to treat to those just below, which can provide an estimate of the causal effect of BP management on dementia (Adina Zeki Al Hazzouri, Columbia University Mailman School of Public Health, personal communication, 2020).

## Outcome

Researchers conducting observational studies often evaluate associations with numerous outcome measures. The ability to triangulate across outcomes can be a strength of observational research. This is particularly true for observational research in ADRD. Dementia is a syndrome, characterized by cognitive and functional impairment preceded by cognitive decline (59). It has a long preclinical phase, characterized by accelerated cognitive decline and accumulation of brain pathology (60). Although brain pathologies are used to define etiologic subtypes (61–63), the correspondence between brain pathology and dementia is imperfect (62). Thus, examination of associations with multiple measures of brain and cognitive health can provide reassurance

that associations are not attributable to outcome-specific sources of bias, allows study of the preclinical phase, and can provide insight into etiologic mechanisms. However, the wide variety of measures available in the observational study setting can also be a limitation. Identical measures are rarely used in studies, reporting is often selective, and it can be difficult to understand whether heterogeneity of findings is related to heterogeneity of measurement or other factors. For example, observational studies of risk factors for ADRD and related outcomes rarely use the same cognitive test batteries, dementia ascertainment approaches, or even the same cutoffs for classifying amyloid positron emission tomography scans (64, 65), and use of different measures of the same construct has been shown to affect study findings (66–68).

To the contrary, RCTs must specify a primary outcome—often with considerations related to Food and Drug Administration approval in mind—and trial findings are evaluated on the basis of findings with this outcome. For example, in most trials testing pharmacologic BP lowering for ADRD prevention, researchers use cognitive impairment, dementia, or significant cognitive decline as their outcome (Web Table 1). There is a strong rationale behind this choice. These are the outcomes that matter to the patient. However, these outcomes are not ideal, because they represent the end state of a long and variable pathologic process and are unlikely to occur during the follow-up period of the trial unless the underlying pathologic process has already begun; thus, most trials are not able to test effects on disease initiation, only disease progression. Similarly, these outcomes are common to multiple etiologic processes. If treatment affects only 1 of these etiologies, the impact of the intervention on risk will be modest at best (61). Although use of a causal intermediate outcome would be preferable, our poor understanding of the disease processes that lead to ADRD currently precludes use of biomarkers of pathologic change as outcomes in RCTs. Likewise, variation in cognitive test batteries across studies contributes to difficulties in identifying robust measures for use in the trial setting, and cognitive change does not necessarily reflect dementia pathogenesis.

Observational studies can better inform clinical trials through robust and systematic reporting of estimated effects with all available outcomes. Reporting of this nature will allow assessment of the likely heterogeneity of the strength of effect with different measures of related outcomes, which will contribute to selection of RCT outcome measures. In the context of outcomes like dementia, where impact of interventions may take years to decades to detect, observational studies will necessarily be instrumental in identifying and validating intermediate outcomes that can then be used as outcomes in both observational studies and RCTs (69).

## Analysis

Observational studies are used to test multiple hypotheses. Although an observational analysis may be motivated by a specific hypothesis, analyses are rarely specified in advance, and exploratory data analysis often informs final reporting. Ultimately, observational studies typically report conditional contrasts of risk across exposure groups, using

**Table 2.** Archetypal Questions Asked and Answered by Observational Studies and Randomized Controlled Trials in the Context of Blood Pressure and Alzheimer Disease or Related Dementias Prevention

Observational Study	RCT
Of people from community X, do those with a SBP of 140 mm Hg at age 50 years have lower dementia risk at age 70 years than those with a SBP of 120 mm Hg at age 50 years, conditional on sociodemographic characteristics, health, and health behaviors at age 50 years?	In persons we can recruit at our clinics, who are at high risk of cardiovascular disease, who are likely to adhere to the study protocol, and have a SBP >160 mm Hg despite available of clinical care and current standards of care, does being randomly assigned to receive Y dose of drug Z for up to 2 years reduce dementia risk compared with being randomly assigned to receive placebo for up to 2 years?

Abbreviations: ADRD, Alzheimer disease and related dementias; RCT, randomized controlled trial; SBP, systolic blood pressure.

regression adjustment to address confounding. The stated or unstated goal of such studies is often to estimate the average treatment effect (ATE); that is, how much would the outcome differ on average if each person were exposed compared with if they were unexposed? To the contrary, RCTs are generally designed to test a single primary hypothesis and a set of predefined secondary hypotheses. Analysis plans are prespecified and focus on ITT estimates, which preserve the benefits of randomization on risk of bias (e.g., balance of measured and unmeasured factors eliminates bias due to confounding). However, the ITT provides the marginal average treatment effect of randomization, rather than treatment.

Observational studies can be used to estimate effects that align with those quantified in RCTs. For example, we can design observational analyses to estimate the ATE for an intervention that is plausible to randomly assign in a trial (e.g., the effect of initiating of antihypertensive medication use on incident dementia). Although such analyses are particularly useful, they will not produce a parameter that would equal the ITT estimate from the corresponding clinical trial. Observational studies necessarily estimate the average treatment effect of treatment, not randomization. Because correspondence between the randomization and treatment received is imperfect, and there is often measurement error in observational study data, we would not expect the ITT estimate to match the ATE. That said, the ATE may provide a plausible upward bounds of a potential ITT effect. It may also be useful to use the ATE to estimate the likely ITT effect in a trial, under assumptions about the degree of attenuation due to nonadherence and the influence of measurement error, or as an estimate of the per-protocol effect that may be observed in the context of a trial.

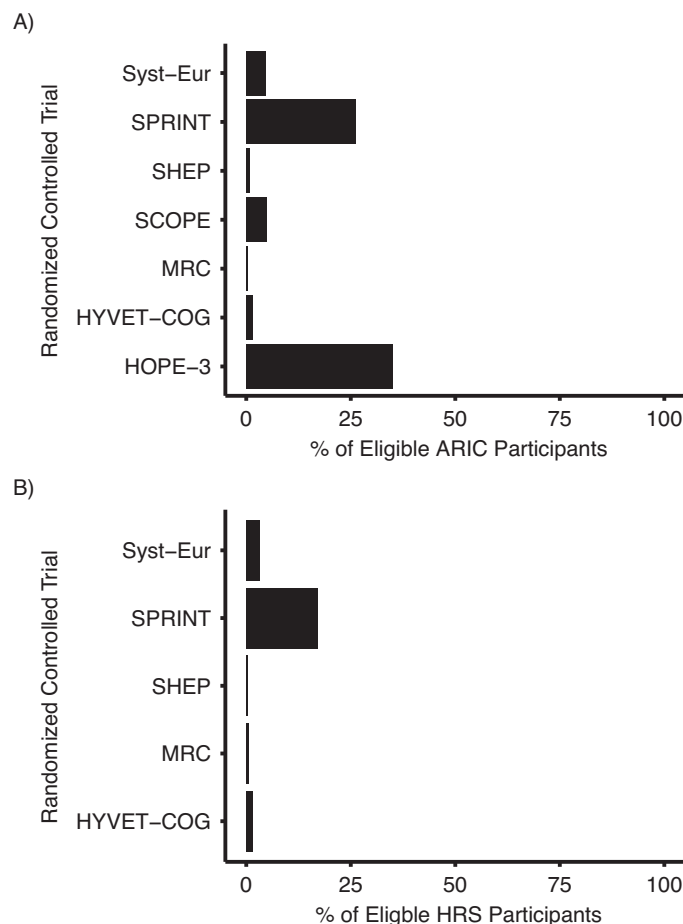
We note that modern causal inference methods may be particularly useful when attempting to estimate the ATE or other causal contrasts relevant to RCT design (e.g., marginal structural models or structural nested models for the effects of time-varying treatment (70, 71); causal mediation methods, particularly interventional direct/indirect effects for the effects of intervention on causal mediators (72); the estimation of the effect of treatment on statistical surrogates as a proxy for long-term effects of treatment (73)). All of these methods come with their own assumptions, strengths, and limitations, and we refer the reader to the extensive literature on these methods.

## EMULATION OF RCTS IN OBSERVATIONAL DATA

Currently, researchers conducting observational studies are often asking and answering questions of little relevance to RCT design (Table 2). Earlier we suggested ways to adjust observational data analyses to make observational analyses better inform clinical trial design. However, in isolation, implementation of these suggestions one by one will only marginally improve translation of observational studies to clinical trials.

The questions asked and answered in observational data do not have to differ substantially from those asked and answered by RCTs. We can use observational data to analyze the observational data in a way that explicitly attempts to answer questions that could be tested in an RCT (i.e., we can emulate the target trial in the observational data) (74, 75). For example, we could use observational data to answer the question of whether, among those with an indication for BP management, does initiating antihypertensive medication use reduce dementia risk by at least 10% over the next 5 years. In fact, many of the ideas we discuss above are pieces of the overall approach one would take emulate a target trial in observational data.

At its core, target trial emulation is simply an organizational structure for causal inference. It focuses the researcher on defining the eligibility criteria, treatment strategies, assignment procedures, follow-up period, outcome, causal contrast of interest, and analysis plan for observational analyses in such a way that it corresponds to the desired target trial. Although there are limits on what sorts of trials we can realistically emulate (74), and data availability can be a substantial barrier to implementation, target trial emulation is possible. For example, a recent study using data from the Rotterdam Study estimated the effect of hypothetical interventions on SBP and smoking on risk stroke and dementia (36). Results suggest that SBP reductions would reduce risk of stroke, but not dementia, over a 15-year intervention period. This is consistent with the findings of existing RCTs, which show robust effects of pharmacologic interventions designed to lower BP on cardiovascular health but not dementia (Web Table 1). Analyses of this nature could be used to explore a variety of related questions that are useful to RCT design, including, “Is possible to observe a beneficial effect of treatment within the ethical, safety, and



**Figure 1.** Proportion of Atherosclerosis Risk in Communities (ARIC) and Health and Retirement Study (HRS) participants who are and are not in the target population for selected randomized controlled trials (RCTs) of blood pressure management for Alzheimer disease and related dementias (ADRD) prevention. (A) The proportion of ARIC visit-5 participants. (B) Proportion of HRS wave 2006 or 2008 participants who participated in the enhanced face-to-face interview, allowing collection of biomarker and physical measurement data, who are in the target population of selected RCTs of blood pressure management for the prevention of ADRD. Selected RCTs do not require presence of cardiovascular disease, diabetes, or prior stroke for eligibility. Target populations are defined as presented in Table 1. Note: We were unable to estimate proportions in HRS for the Study on Cognition and Prognosis in the Elderly (SCOPE) or Heart Outcomes Prevention Evaluation-3 (HOPE-3), because of the absence of data required for determining target population eligibility; we excluded all persons who were taking antihypertensive medications when evaluating Systolic Hypertension in the Elderly Program (SHEP) criteria (SHEP allowed participation of persons taking antihypertensive medications if they met systolic blood pressure criteria after withdrawal of medication use); we assumed HRS participants younger than 70 years do not have dementia (dementia status is algorithmically defined only for those older than 70 years); and all estimates reflect the proportion in the target population among those with nonmissing data on variables needed to define target population membership. Abbreviations: HYVET-COG, Hypertension in the Very Elderly Trial Cognitive Function Assessment; MRC, Medical Research Council; SPRINT, Systolic Blood Pressure Intervention Trial; SYST-EUR, Systolic Hypertension in Europe.

feasibility constraints of an RCT?” and “What would the trial with the greatest likelihood of providing a definitive answer look like?”

#### ISSUES OF DATA AVAILABILITY AND POTENTIAL SOLUTIONS

One pervasive difficulty with using observational data to better inform RCT design is data availability. Although this is particularly true for target trial emulation, data availability can be a barrier to implementation of many of our recom-

mendations (Table 1). Only a small fraction of participants in observational studies would be eligible for inclusion in most RCTs. For example, the proportion of ARIC or Health and Retirement Study participants who are part of the target population of selected RCTs of BP lowering for ADRD prevention is often very small (Figure 1). The schedule of data collection in observational studies is also often limited compared with the frequency and depth of relevant data collected by corresponding trials. Most cohort studies collect data at intervals of years, whereas trials typically collect data at shorter intervals. Although administrative databases,



such as electronic health records or claims databases (e.g., Medicare) offer an attractive alternative, these sources have limitations. Availability of data is uniquely tied to participant characteristics, because the type and timing of data are determined by timing of patient encounters and clinical or billing needs. In addition, privacy concerns about this type of data may preclude pooling or linkage of data with other sources.

Observational data also frequently have limited diversity of study participants. Historically, both observational studies and RCTs in the United States recruited samples that were predominantly non-Hispanic White. Although new RCTs are increasingly diverse, the diversity of existing observational studies is often limited by historical decisions about recruitment. These studies often lack sufficient representation from racial/ethnic minority groups to robustly estimate effects in these groups (76). This is problematic because risk is unequally distributed, and treatment effects may differ across groups. For example, Black Americans are about twice as likely to have cognitive impairment or dementia than Whites (77–81), and drug responses varied by race/ethnicity in approximately 20% of new molecular entities tested between 2008 and 2013 (82).

In this context, a combination of data sources is often a prerequisite to implementing methods that increase relevance of observational studies to trial design (Table 1) or for target trial emulation. Assuming that all studies pull from the same source population, data from multiple studies can be combined using a variety of existing approaches. When individual data are available, it can be pooled and harmonized. Although seemingly simple, in practice harmonization can be difficult or impossible because of differences in measures collected and measurement schedules across studies. Collection of a consistent set of core measures across observational studies (and their related RCTs), supplemented with study-specific measures, would help overcome this difficulty and would allow for better cross-cohort harmonization.

When data are comparable enough to allow synthesis but cannot be directly pooled and harmonized (e.g., due to privacy restrictions), other approaches may be necessary. Meta-analysis (83), related Bayesian approaches (84, 85), coordinated analyses (86, 87), or aggregate-data based approaches (88) are often used to produce combined effect estimates. Partially or fully synthetic data approaches may also permit data sharing and pooled analyses while remaining consistent with data privacy goals (89–97).

Although convenient, the assumption that different studies pull from the same source population underlies the validity of all of these methods for combining data sources. However, it should not be assumed that this assumption holds (68). Additional methods work (e.g., within the transportability framework (42–45)) will be necessary to allow combination of data sources where this assumption is violated.

## CONCLUSIONS

Well-powered RCTs are considered the gold standard for causal inference in biomedical research (98). In a perfect world, there would be a seamless handoff from obser-

vatational researchers to trialists. Observational researchers would build the case for whether and how to intervene on a particular risk factor and provide insight into choice of target population, outcome, and follow-up time. Trialists would then use that detailed evidence to make decisions about whether and how to field an RCT. Unfortunately, the questions asked and answered by observational studies often have little relevance to the decisions that need to be made when designing and fielding an RCT. However, observational studies can be analyzed in ways that make them more relevant to trial design, and better use of observational studies to inform choices around clinical trial design should increase the likelihood of successful, or at least definitive trials. Ultimately, analyses in observational studies should inform not only what the RCT design should be but also—given the substantial cost in money, time, and public trust accompanying nondefinitive results—whether an RCT should be fielded at all.

## ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Milken Institute School of Public Health, George Washington University, Washington, DC, United States (Melinda C. Power, Brittany Engelman, Jingkai Wei); and Department of Epidemiology and Biostatistics, School of Medicine, University of California San Francisco, San Francisco, California, United States (M. Maria Glymour).

This work was funded by the National Institute of Aging (NIA; grant R01AG057869 to M.C.P. and M.M.G.). The Atherosclerosis Risk in Communities (ARIC) Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute (NHLBI) contracts (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700005I, and HHSN268201700004I). NHLBI grants U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, and 2U01HL096917 (cofunded by the National Institute of Neurological Disorders and Stroke, NIA, and National Institute on Deafness and Other Communication Disorders) supported the collection of neurocognition data, with previous brain magnetic resonance imaging examinations funded by an NHLBI (grant R01-HL70825).

The ARIC study has established research access policies; requests will be promptly reviewed for confidentiality or intellectual property restrictions and will not unreasonably be refused. Individual-level patient data may further be restricted by consent, confidentiality, or privacy laws/considerations. For more information on these policies and to request ARIC data, please visit <https://sites.cscs.unc.edu/aric/distribution-agreements>. The HRS also has established procedures by which interested investigators can obtain data from the study. Information on the related policies and procedures can be found on the HRS website: <https://hrs.isr.umich.edu/>

The authors thank the staff and participants of the ARIC study for their important contributions and Victoria Willens and Erin Bennett for their contributions to this work.

The views expressed in this article are those of the authors and do not reflect those of the NIH.

Conflict of interest: The authors report having received funding from the US NIH, the US Department of Defense, and DC Health.

## REFERENCES

- Prince M, Ali GC, Guerchet M, et al. Recent global trends in the prevalence and incidence of dementia, and survival with dementia. *Alzheimers Res Ther.* 2016;8(1):23.
- Prince M, Bryce R, Albanese E, et al. The global prevalence of dementia: a systematic review and meta-analysis. *Alzheimers Dement.* 2013;9(1):63–75.e2.
- Alzheimer's Association. Alzheimer's disease facts and figures special report: ethnicity and Alzheimer's in America. 2021. <https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf>. Accessed March 13, 2021.
- El-Hayek YH, Wiley RE, Khoury CP, et al. Tip of the iceberg: assessing the global socioeconomic costs of Alzheimer's disease and related dementias and strategic implications for stakeholders. *J Alzheimers Dis.* 2019;70(2):323–341.
- Hurd MD, Martorell P, Delavande A, et al. Monetary costs of dementia in the United States. *N Engl J Med.* 2013;368(14):1326–1334.
- Wimo A, Guerchet M, Ali G-C, et al. The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimers Dement.* 2017;13(1):1–7.
- Hebert LE, Weuve J, Scherr PA, et al. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. *Neurology.* 2013;80(13):1778–1783.
- Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers Res Ther.* 2014;6(4):37.
- Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *Lancet.* 2017;390(10113):2673–2734.
- Norton S, Matthews FE, Barnes DE, et al. Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. *Lancet Neurol.* 2014;13(8):788–794.
- Mukadam N, Anderson R, Knapp M, et al. Effective interventions for potentially modifiable risk factors for late-onset dementia: a costs and cost-effectiveness modelling study. *The Lancet Healthy Longev.* 2020;1(1):e13–e20.
- Brookmeyer R, Abdalla N, Kawas CH, et al. Forecasting the prevalence of preclinical and clinical Alzheimer's disease in the United States. *Alzheimers Dement.* 2018;14(2):121–129.
- Wu YT, Fratiglioni L, Matthews FE, et al. Dementia in western Europe: epidemiological evidence and implications for policy making. *Lancet Neurol.* 2016;15(1):116–124.
- Andrieu S, Coley N, Lovestone S, et al. Prevention of sporadic Alzheimer's disease: lessons learned from clinical trials and future directions. *Lancet Neurol.* 2015;14(9):926–944.
- Qiu C, Winblad B, Fratiglioni L. The age-dependent relation of blood pressure to cognitive function and dementia. *Lancet Neurol.* 2005;4(8):487–499.
- Walker KA, Power MC, Gottesman RF. Defining the relationship between hypertension, cognitive decline, and dementia: a review. *Curr Hypertens Rep.* 2017;19(3):24.
- Ou YN, Tan CC, Shen XN, et al. Blood pressure and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 209 prospective studies. *Hypertension.* 2020;76(1):217–225.
- Wahidi N, Lerner AJ. Blood pressure control and protection of the aging brain. *Neurotherapeutics.* 2019;16(3):569–579.
- Power MC, Tchetgen Tchetgen EJ, Sparrow D, et al. Blood pressure and cognition: factors that may account for their inconsistent association. *Epidemiology.* 2013;24(6):886–893.
- Power MC, Schneider AL, Wruck L, et al. Life-course blood pressure in relation to brain volumes. *Alzheimers Dement.* 2016;12(8):890–899.
- Walker KA, Sharrett AR, Wu A, et al. Association of midlife to late-life blood pressure patterns with incident dementia. *JAMA.* 2019;322(6):535–545.
- SPRINT MIND Investigators for the SPRINT Research Group. Effect of intensive vs standard blood pressure control on probable dementia: a randomized clinical trial. *JAMA.* 2019;321(6):553–561.
- The SPRINT Research Group. A randomized trial of intensive versus standard blood-pressure control. *New Eng J Med.* 2015;373(22):2103–2116.
- Hughes D, Judge C, Murphy R, et al. Association of blood pressure lowering with incident dementia or cognitive impairment: a systematic review and meta-analysis. *JAMA.* 2020;323(19):1934–1944.
- Panagiotou OA, Jaljuli I, Heller R. Replicability of treatment effect in study of blood pressure lowering with dementia. *JAMA.* 2020;324(14):1465–1466.
- McGuinness B, Todd S, Passmore P, et al. The effects of blood pressure lowering on development of cognitive impairment and dementia in patients without apparent prior cerebrovascular disease. *Cochrane Database Syst Rev.* 2006;(2):CD004034.
- van Middelaar T, van Vught LA, van Gool WA, et al. Blood pressure-lowering interventions to prevent dementia: a systematic review and meta-analysis. *J Hypertens.* 2018;36(9):1780–1787.
- Williamson JD, Supiano MA, Applegate WB, et al. Intensive vs standard blood pressure control and cardiovascular disease outcomes in adults aged  $\geq 75$  years: a randomized clinical trial. *JAMA.* 2016;315(24):2673–2682.
- US Federal Government. Award search. <https://www.usaspending.gov/search>. Accessed May 11, 2021.
- Gottesman RF, Albert MS, Alonso A, et al. Associations between midlife vascular risk factors and 25-year incident dementia in the Atherosclerosis Risk in Communities (ARIC) cohort. *JAMA Neurol.* 2017;74(10):1246–1254.
- Gottesman RF, Schneider AL, Albert M, et al. Midlife hypertension and 20-year cognitive change: the Atherosclerosis Risk in Communities Neurocognitive Study. *JAMA Neurol.* 2014;71(10):1218–1227.
- Power MC, Tingle JV, Reid RI, et al. Midlife and late-life vascular risk factors and white matter microstructural integrity: the Atherosclerosis Risk in Communities Neurocognitive Study. *J Am Heart Assoc.* 2017;6(5):e005608.
- Ruitenberg A, Skoog I, Ott A, et al. Blood pressure and risk of dementia: results from the Rotterdam study and the Gothenburg H-70 study. *Dement Geriatr Cogn Disord.* 2001;12(1):33–39.
- in't Veld BA, Ruitenberg A, Hofman A, et al. Antihypertensive drugs and incidence of dementia: the Rotterdam study. *Neurobiol Aging.* 2001;22(3):407–412.
- Ma Y, Wolters FJ, Chibnik LB, et al. Variation in blood pressure and long-term risk of dementia: a population-based cohort study. *PLoS Med.* 2019;16(11):e1002933.

36. Rojas-Saunero LP, Hilal S, Murray EJ, et al. Hypothetical blood-pressure-lowering interventions and risk of stroke and dementia. *Eur J Epidemiol.* 2021;36(1):69–79.
37. The Atherosclerosis Risk in Communities Study. Atherosclerosis Risk in Communities Study Protocol Manual 2: Cohort Component Procedures. 1988. [https://sites.csc.unc.edu/aric/sites/default/files/public/manuals/Cohort\\_Procedures.1\\_2.pdf](https://sites.csc.unc.edu/aric/sites/default/files/public/manuals/Cohort_Procedures.1_2.pdf). Accessed June 18, 2020.
38. Hofman A, Grobbee DE, de Jong PT, et al. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol.* 1991;7(4):403–422.
39. Casey JA, Schwartz BS, Stewart WF, et al. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health.* 2016;37:61–81.
40. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol.* 2007;17(9):643–653.
41. Alonso A, Seguí-Gómez M, Irala J, et al. Predictors of follow-up and assessment of selection bias from dropouts using inverse probability weighting in a cohort of university graduates. *Eur J Epidemiol.* 2006;21(5):351–358.
42. Westreich D, Edwards JK, Lesko CR, et al. Transportability of trial results using inverse odds of sampling weights. *Am J Epidemiol.* 2017;186(8):1010–1014.
43. Stuart EA, Ackerman B, Westreich D. Generalizability of randomized trial results to target populations: design and analysis possibilities. *Res Soc Work Pract.* 2018;28(5):532–537.
44. Dahabreh IJ, Robertson SE, Steingrimsdóttir JA, et al. Extending inferences from a randomized trial to a new target population. *Stat Med.* 2020;39(14):1999–2014.
45. Bareinboim E, Pearl J. A general algorithm for deciding transportability of experimental results. *J Causal Inference.* 2013;1(1):107–134.
46. Kivipelto M, Helkala EL, Laakso MP, et al. Apolipoprotein E epsilon4 allele, elevated midlife total cholesterol level, and high midlife systolic blood pressure are independent risk factors for late-life Alzheimer disease. *Ann Intern Med.* 2002;137(3):149–155.
47. Applegate WB, Pressel S, Wittes J, et al. Impact of the treatment of isolated systolic hypertension on behavioral variables. Results from the systolic hypertension in the elderly program. *Arch Intern Med.* 1994;154(19):2154–2160.
48. Peters R, Beckett N, Forette F, et al. Incident dementia and blood pressure lowering in the hypertension in the very elderly trial cognitive function assessment (HYVET-COG): a double-blind, placebo controlled trial. *Lancet Neurol.* 2008;7(8):683–689.
49. Murray AM, Hsu FC, Williamson JD, et al. ACCORDION MIND: results of the observational extension of the ACCORD MIND randomised trial. *Diabetologia.* 2017;60(1):69–80.
50. Alzheimer's Association. Alzheimer's association funds 2-year extension of SPRINT MIND study. 2019. <https://www.alz.org/news/2019/alzheimer-s-association-funds-two-year-extension-o>. Access date: March 8, 2022.
51. Wang M, Liao X, Laden F, et al. Quantifying risk over the life course - latency, age-related susceptibility, and other time-varying exposure metrics. *Stat Med.* 2016;35(13):2283–2295.
52. Lithell H, Hansson L, Skoog I, et al. The Study on Cognition and Prognosis in the Elderly (SCOPE): principal results of a randomized double-blind intervention trial. *J Hypertens.* 2003;21(5):875–886.
53. Kennelly SP, Lawlor BA, Kenny RA. Blood pressure and dementia - a comprehensive review. *Ther Adv Neurol Disord.* 2009;2(4):241–260.
54. Kaiser P, Arnold AM, Benkeser D, et al. Comparing methods to address bias in observational data: statin use and cardiovascular events in a US cohort. *Int J Epidemiol.* 2018;47(1):246–254.
55. Seeger JD, Williams PL, Walker AM. An application of propensity score matching using claims data. *Pharmacoepidemiol Drug Saf.* 2005;14(7):465–476.
56. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study Design in Pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep.* 2015;2(4):221–228.
57. Petersen ML, Porter KE, Gruber S, et al. Diagnosing and responding to violations in the positivity assumption. *Stat Methods Med Res.* 2012;21(1):31–54.
58. Ogilvie D, Adams J, Bauman A, et al. Using natural experimental studies to guide public health action: turning the evidence-based medicine paradigm on its head. *J Epidemiol Community Health.* 2020;74(2):203–208.
59. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V)*. Arlington, VA: American Psychiatric Association; 2013.
60. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7(3):280–292.
61. Power MC, Mormino E, Soldan A, et al. Combined neuropathological pathways account for age-related risk of dementia. *Ann Neurol.* 2018;84(1):10–22.
62. Schneider JA, Arvanitakis Z, Bang W, et al. Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology.* 2007;69(24):2197–2204.
63. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7(3):263–269.
64. Gottesman RF, Schneider AL, Zhou Y, et al. The ARIC-PET amyloid imaging study: brain amyloid differences by age, race, sex, and APOE. *Neurology.* 2016;87(5):473–480.
65. Landau S, Jagust W. Florbetapir processing methods. 2015. [https://adni.bitbucket.io/reference/docs/UCBERKELEYAV45/ADNI\\_AV45\\_Methods\\_JagustLab\\_06.25.15.pdf](https://adni.bitbucket.io/reference/docs/UCBERKELEYAV45/ADNI_AV45_Methods_JagustLab_06.25.15.pdf). Accessed March 8, 2022.
66. Power MC, Gianattasio KZ, Ciarleglio A. Implications of the use of algorithmic diagnoses or Medicare claims to ascertain dementia. *Neuroepidemiology.* 2020;54(6):462–471.
67. Brookmeyer R, Evans DA, Hebert L, et al. National estimates of the prevalence of Alzheimer's disease in the United States. *Alzheimers Dement.* 2011;7(1):61–73.
68. Gianattasio KZ, Bennett EE, Wei J, et al. Generalizability of findings from a clinical sample to a community-based sample: a comparison of ADNI and ARIC. *Alzheimers Dement.* 2021;17(8):1265–1276.
69. DeMets DL, Psaty BM, Fleming TR. When can intermediate outcomes be used as surrogate outcomes? *JAMA.* 2020;323(12):1184–1185.
70. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology.* 2000;11(5):561–570.
71. Vansteelandt S, Joffe M. Structural nested models and G-estimation: the partially realized promise. *Stat Sci.* 2014;29(4):707–731.

72. Nguyen TQ, Schmid I, Stuart EA. Clarifying causal mediation analysis for the applied researcher: defining effects based on what we want to learn. *Psychol Methods*. 2021; 26(2):255–271.
73. Athey S, Chetty R, Imbens GW, et al. Estimating treatment effects using multiple surrogates: the role of the surrogate score and the surrogate index [preprint; revised February 2020]. arXiv.org. (<https://ideas.repec.org/p/arx/papers/1603.09326.html>). Accessed March 8, 2022.
74. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758–764.
75. Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. *Eur J Epidemiol*. 2017;32(6): 473–475.
76. Shin J, Doraiswamy PM. Underrepresentation of African-Americans in Alzheimer’s trials: a call for affirmative action. *Front Aging Neurosci*. 2016;8:123.
77. Tang MX, Cross P, Andrews H, et al. Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology*. 2001;56(1):49–56.
78. Power MC, Bennett EE, Turner RW, et al. Trends in relative incidence and prevalence of dementia across non-Hispanic black and white individuals in the United States, 2000–2016. *JAMA Neurol*. 2021;78(3):275–284.
79. Fitzpatrick AL, Kuller LH, Ives DG, et al. Incidence and prevalence of dementia in the cardiovascular health study. *J Am Geriatr Soc*. 2004;52(2):195–204.
80. Rajan KB, Weuve J, Barnes LL, et al. Prevalence and incidence of clinically diagnosed Alzheimer’s disease dementia from 1994 to 2012 in a population study. *Alzheimers Dement*. 2019;15(1):1–7.
81. Steenland K, Goldstein FC, Levey A, et al. A meta-analysis of Alzheimer’s disease incidence and prevalence comparing African-Americans and Caucasians. *J Alzheimers Dis*. 2016; 50(1):71–76.
82. Ramamoorthy A, Pacanowski MA, Bull J, et al. Racial/ethnic differences in drug disposition and response: review of recently approved drugs. *Clin Pharmacol Ther*. 2015;97(3): 263–273.
83. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–188.
84. Ades AE, Sutton AJ. Multiparameter evidence synthesis in epidemiology and medical decision-making: current approaches. *J R Stat Soc A Stat Soc*. 2006;169(1):5–35.
85. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat Med*. 2003;22(23): 3687–3709.
86. Zammit AR, Piccinin AM, Duggan EC, et al. A coordinated multi-study analysis of the longitudinal association between handgrip strength and cognitive function in older adults. *J Gerontol B Psychol Sci Soc Sci*. 2019;76(2): 229–241.
87. Hofer SM, Piccinin AM. Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychol Methods*. 2009; 14(2):150–164.
88. Li X, Fireman BH, Curtis JR, et al. Validity of privacy-protecting analytical methods that use only aggregate-level information to conduct multivariable-adjusted analysis in distributed data networks. *Am J Epidemiol*. 2019; 188(4):709–723.
89. Rubin DB. Statistical disclosure limitation. *J Off Stat*. 1993; 9(2):461–468.
90. Little R. Statistical analysis of masked data. *J Off Stat*. 1993; 9(2):407–426.
91. Reiter JP. Using CART to generate partially synthetic public use microdata. *J Off Stat*. 2005;21(2):441.
92. Drechsler J. *Using Support Vector Machines for Generating Synthetic Datasets*. Berlin, Germany: Springer Berlin Heidelberg; 2010:148–161.
93. Caiola G, Reiter JP. Random forests for generating partially synthetic, categorical data. *Trans Data Privacy*. 2010;3(1): 27–42.
94. Dandekar A, Zen RAM, Bressan S. *A Comparative Study of Synthetic Dataset Generation Techniques*. Cham: Springer International Publishing; 2018:387–395.
95. Young J, Graham P, Penny R. Using Bayesian networks to create synthetic data. *J Off Stat*. 2009;25(4):549.
96. Goncalves A, Ray P, Soper B, et al. Generation and evaluation of synthetic patient data. *BMC Med Res Methodol*. 2020;20(1):108.
97. Filshtein TJ, Li X, Zimmerman SC, et al. Proof of concept example for use of simulation to allow data pooling despite privacy restrictions. *Epidemiology*. 2021;32(5): 638–647.
98. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25):1887–1892.