

**National Institute on Aging (NIA)
Division of Behavioral and Social Research (BSR)**

Non-pharmacological Approaches to the Early Prevention of AD/ADRD

MAY 24-25, 2021

September 28, 2021



This meeting summary was prepared by Rose Li & Associates, Inc., under contract to the National Institute on Aging (NIA) Division of Behavioral and Social Research (BSR). The views expressed in this document reflect both individual and collective opinions of the meeting participants and not necessarily those of NIA. Contributions to this summary by the following individuals are gratefully acknowledged: Caroline Sferrazza, Jessica Boten, Adam M. Brickman, Dana Carluccio, Linda M. Collins, Emilie Courtin, Jason Doctor, Richard Gershon, Maria Glymour, Jonathan W. King, Beth Mormino, Lauren Nicholas, Melinda Power, Yakeel Quiroz, Martin Sliwinski, Bethany Stokes, Sonja Swanson, Nancy Tuveson, Rachel Whitmer, and Julie Zissimopoulos.

Table of Contents

Acronyms	iii
Executive Summary	1
Meeting Summary	3
Welcome Remarks	3
State of the Science and Précis for the Meeting	3
Panel 1: Assessing Target Engagement.....	4
Imaging Biomarkers That Could Index Change	4
Blood-based Biomarkers: Opportunities for Early Prevention Trials of AD/ADRD	5
Measuring Subtle Change in Psychological Processes: Challenges and Opportunities	6
Panel Discussion	7
Panel 2: Measuring Outcomes	8
Standard Neuropsychology as an Outcome: Challenges and Opportunities	8
Remote Cognitive Measurement as an Outcome	9
Found Measures of Cognition	10
Panel Discussion	11
Panel 3: Issues in Intervention Development: How to Design the Most Informative Trials.....	12
Developing Effective, Affordable, Scalable, and Efficient Interventions.....	12
Who Do We Really End Up Recruiting into Our Trials and Research?	13
Large Trials in the Health Care System to Detect and Prevent Dementia	14
Panel Discussion	15
Panel 4: Non-intervention Development: Generating Causal Evidence from Observational and Non-clinical Studies to Inform the Prevention Research Agenda	16
Causal Inference from Observational Data	16
Quasi-Experimental Methods	17
Measuring Effects of Treatment Variation in Healthcare Systems and Their Effects on Cognition ..	19
Chronic Conditions and Risk of AD/ADRD: Advancing Prevention Analyses to Inform Population Health	20
Panel Discussion	21
General Discussion.....	22
Appendix A: Meeting Agenda	24

Acronyms

A4	Anti-Amyloid Treatment in Asymptomatic Alzheimer's
A β	β -amyloid
ACEI	ACE inhibitor
ACTIVE	Advanced Cognitive Training for Independent and Vital Elderly
AD	Alzheimer's disease
ADAD	autosomal dominant Alzheimer's disease
ADAMS	Aging, Demographics and Memory Study
ADNI	Alzheimer's Disease Neuroimaging Initiative
ADRD	Alzheimer's disease and related dementias
AHT	anti-hypertensive
API	Alzheimer's Prevention Initiative
ARB	angiotensin receptor blocker
ARIC	Atherosclerosis Risk in Communities Study
BBB	blood-brain barrier
BSR	Division of Behavioral and Social Research
CAIDE	Cardiovascular Risk Factors, Aging, and Incidence of Dementia
CSF	cerebrospinal fluid
DN	Division of Neuroscience
<i>EASE</i>	effectiveness against affordability, scalability, and efficiency
EHR	electronic health record
EMA	ecological momentary assessment
FFS	fee-for-service
FNAME	Face Name Associative Memory Exam
HRS	Health and Retirement Study
MarkVCID	Biomarkers for Vascular Contributions to Cognitive Impairment and Dementia
MTB	Mobile Toolbox
NASEM	National Academies of Sciences, Engineering, and Medicine
NIA	National Institute on Aging
PACC	Preclinical Alzheimer's Cognitive Composite
PET	positron emission tomography
PmNTB	POINTER modified Neuropsychological Test Battery
PRO	patient-reported outcome
PROMIS	Patient Reported Outcomes Measurement Information System
pTau	phosphorylated tau
RAS	renin-angiotensin system
RCT	randomized controlled trial
RNT	repetitive negative thinking
RWE	real world evidence
SMART	sequential multiple assignment randomized trial
SPRINT	Systolic Blood Pressure Intervention Trial

Executive Summary

On May 24-25, 2021, the National Institute on Aging (NIA) Division of Behavioral and Social Research (BSR) held a workshop to review non-pharmacological approaches for early prevention of Alzheimer's disease and Alzheimer's disease related dementias (AD/ADRD). Although preventive non-pharmacological interventions beginning in midlife offer an opportunity to eliminate disease burden without the risks or costs associated with chronic drug regimens, a traditional randomized controlled trial (RCT) with decades-long follow-up would face immediate and significant design challenges. The goal of this workshop was to address the core methodological challenges associated with developing and accumulating evidence to support preventive interventions for AD/ADRD with respect to the populations, targets, and outcomes that should be the focus of the prevention research agenda. The workshop was composed of four panels that each addressed a different aspect of intervention development: (1) identifying biomarkers that would show early evidence of disease modification and possibly offer surrogate outcomes, (2) measuring cognitive outcomes using instruments and methods consistent with large-scale trials of diverse populations, (3) designing the most informative trials, and (4) generating causal evidence from observational and non-clinical studies to either mirror RCTs or guide future RCTs.

The first panel offered an introduction to potentially modifiable molecular, neural, and psychological processes early in the trajectory of AD/ADRD pathology that may serve as biomarkers or surrogate outcomes for preventive interventions in early or midlife and facilitate evaluation of new interventions. This panel focused on the suitability of currently available measures for use in prevention trials and the evidence needed to validate each type of measure (e.g., reliable linkage to functional outcomes, validity across diverse groups and life stages).

The second panel focused on the advantages and disadvantages of different types of cognitive and dementia-relevant outcome measures (e.g., neuroimaging measures) that have immediate relevance to patients or can be used to validate potential targets. Among the cognitive measures considered were standard neuropsychological assessments, remote cognitive measurements feasible for large population-based samples, and "found" measures of cognitive status or subtle symptoms of cognitive impairment (e.g., administrative data on daily functioning). Panelists addressed the need for measures that (1) can detect subtle changes in cognition and behavior in early and midlife; (2) can be measured in large and diverse samples, potentially with passive data accrual; and (3) are known to be linked to the distal outcome of AD/ADRD.

During the third panel, the panelists explored how researchers can generate causal evidence through thoughtful study design before launching large-scale prevention trials, thus improving the likelihood of trial success and facilitating the ability to tailor interventions. Panelists emphasized the value of flexible and adaptive experimental strategies (e.g., multi-phase optimization strategy [MOST] designs), the identification of appropriate study populations that represent people affected by AD/ADRD, and the potential to embed prevention trials in large health care systems for prevention research. The electronic health records (EHR) systems of

many large health care systems are a potentially powerful platform for fielding and easily evaluating large-scale preventive or treatment interventions because they both facilitate behavioral nudges and passively collect outcome data.

The fourth panel reviewed advances in causal modeling from observational data, which may improve the ability of observational studies to anticipate effects of proposed interventions by using quasi-experimental approaches to draw rigorous evidence from accidental randomized trials (e.g., arbitrary timing of policy changes). Such approaches can either circumvent the logistical challenges posed by traditional long-term prevention trials beginning in midlife or guide better intervention studies that are more likely to succeed. Panelists discussed the special relevance of health care system data as an opportunity for rigorous research, due to sample size and diversity, prospective longitudinal measurement, and potentially unbiased sources of variation in treatments received. Finally, the panel discussed biologically informed hypothesis testing using real-world evidence tied to dynamic microsimulation, which can be used to anticipate the long-term population impacts on AD/ADRD incidence and prevalence.

The workshop concluded with an overarching discussion of the major themes that emerged over the course of both days. Panelists broadly agreed that emerging data sources and methods offer tremendous opportunity to support AD/ADRD prevention trials in midlife. While such trials are under way, rigorous quasi-experimental evidence can be generated for the interventions most likely to be effective in preventing AD/ADRD. Panelists therefore highlighted the following opportunities in the prevention research agenda for AD/ADRD that could address the field's outstanding needs:

1. Recruitment of larger, more representative samples at all phases of prevention research
2. Creation of a national registry for biospecimens, which could offer a platform for study recruitment; validation of imaging, plasma, or cognitive biomarkers against clinically meaningful outcomes; and passive follow-up for non-invasive measures; and close collaborations with health care systems that have already accrued large volumes of longitudinal data
3. Multidisciplinary research to identify quasi-experimental opportunities, define biologically plausible and specific hypotheses to test in large datasets, promote rigorous causal approaches to evaluating observational data, and help design more efficient and likely-to-succeed RCTs
4. Trials within large health care systems and other structures that passively accrue long-term exposure and outcome data on people affected by AD/ADRD
5. Infrastructure to support measurement development (including biomarkers and cognitive assessments) and trial recruitment

Meeting Summary

Welcome Remarks

Lis Nielsen, PhD, NIA/BSR; Eliezer Masliah, MD, NIA/DN; Maria Glymour, ScD, MS, University of California, San Francisco

The National Institute on Aging (NIA) Division of Behavioral and Social Research (BSR) and Division of Neuroscience (DN) support studies that build an evidence base for a variety of lifestyle and behavioral interventions to promote cognitive and brain health as well as prevent cognitive decline and dementia. Emerging evidence demonstrates that behavioral patterns and psychological profiles observable earlier in life may be associated with risk of or resilience to Alzheimer's disease (AD) and AD-related dementias (ADRD). However, the causal mechanisms that give rise to risk and resilience are not understood. Moreover, BSR and DN support the development of non-pharmacological interventions that improve AD/ADRD-related outcomes, and many of the interventions with the strongest evidence base (e.g., cognitive training, physical activity, blood pressure control) may require behavior changes that are sustained for long time periods. To support understanding of causal risk or resilience mechanisms and promising non-pharmacological interventions, more studies of AD/ADRD are needed that focus on earlier ages and have longer follow-ups.

This workshop features presentations and discussion about the evidence base that is needed to justify a public health agenda focused on long-term lifestyle and behavior changes for the prevention of AD/ADRD. The subject matter experts invited to present at this workshop were tasked with addressing the gaps and opportunities that exist in this field, including the multifaceted challenges of designing more informative AD/ADRD trials in midlife, the difficulty of precisely defining effective interventions, and the conceptualization of meaningful, functional, and mechanistic outcome assessments.

State of the Science and Précis for the Meeting

Jonathan W. King, PhD, NIA

NIA has defined AD/ADRD Research Implementation Milestones as part of the National Plan to Address AD, which aims to prevent and treat AD/ADRD by 2025. The immediate catalyst for this workshop is Milestone 8.D: initiate at least one clinical trial for primary prevention of AD/ADRD among high-risk individuals in midlife, including a non-pharmacological treatment arm. This workshop was also informed by a 2017 report from the National Academies of Sciences, Engineering, and Medicine (NASEM), which suggested numerous ways to construct a stronger evidence base for the prevention of cognitive decline and ADRD. These suggestions included tailoring interventions to individuals at highest risk of cognitive decline and ADRD, initiating more interventions at younger ages with longer follow-ups, using consistent cognitive outcome measures across trials, including biomarkers as intermediate outcomes, increasing participation of underrepresented populations to study intervention effectiveness, and conducting large

trials designed to test the effectiveness of an intervention in broad routine clinical practices or community settings.

NASEM has also highlighted three promising classes of interventions that should be prioritized for AD/ADRD research: blood pressure management for people with hypertension, increased physical activity, and cognitive training. All of these intervention classes are part of NIA's research and clinical trials portfolio, including the Systolic Blood Pressure Intervention Trial (SPRINT) MIND study and the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) trial. However, even when supporting evidence is generated for these interventions, the mechanism of action often remains unclear. For example, the ACTIVE trial demonstrated a 10-year benefit in reasoning and speed of processing from cognitive training, but a combination of factors (e.g., attrition, practice effects, selected outcome measures) prevent the identification of the mechanism for its success.

This workshop focuses on the core aspects of designing trials for non-pharmacological interventions that can elucidate the causal mechanisms of AD/ADRD prevention. The first panel addresses how to measure engagement of a putative target for intervention. The second panel focuses on AD/ADRD-relevant outcomes that can be used to validate putative targets. The third panel explores how to determine whether non-pharmacological interventions are efficacious, including the discrete components of intervention packages (e.g., how interventions are delivered) that may influence efficacy. Finally, the fourth panel introduces ways to gain knowledge about effective interventions outside of traditional trials (e.g., generating causal evidence from observational studies).

Panel 1: Assessing Target Engagement

Chair: Beth Mormino, PhD, Stanford University

Imaging Biomarkers That Could Index Change

Beth Mormino, PhD, Stanford University

Approaches to prevention research may be broadly conceptualized within two categories: secondary prevention, which aims to prevent clinical symptoms in individuals that demonstrate quantitative evidence of pathological changes but lack clinical symptoms, and primary prevention, which aims to prevent the emergence of pathological changes altogether. For example, primary and secondary prevention trials that target amyloid pathology have been designed for individuals who have little evidence of amyloid pathology, such as in the AHEAD 3-45 study (primary prevention, with the goal of preventing future amyloid accumulation), or who have developed amyloid pathology, such as in the Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) study (secondary prevention, with the goal of preventing future cognitive changes). Both primary and secondary intervention strategies for AD/ADRD need to be studied and implemented years or decades before the onset of clinical symptoms in order to influence outcomes because amyloid and tau pathology begin well before cognitive issues appear. The decades between initial amyloid accumulation and clinically meaningful change represent a major barrier for the prevention of AD/ADRD. One way that quantitative biomarkers can

contribute to AD/ADRD intervention trials is by acting as surrogate endpoints, effectively shortening the length of a trial by focusing on measures that will change *before* clinically meaningful change (e.g., progression to mild cognitive impairment or dementia). Quantitative biomarkers that are sensitive to changes in a specific pathway can also be leveraged to measure a hypothesized target or to provide insight into intervention mechanisms by disambiguating changes across multiple pathways. An opportunity to improve the utility of quantitative biomarkers is to develop biomarkers that are sensitive to changes earlier in life. Importantly, more sensitive measurements of cognitive and functional change are needed for biomarker validation.

Amyloid positron emission tomography (PET) is one imaging-based biomarker that may be useful for prevention trials of AD/ADRD. However, even subtle clinically meaningful change can take years to observe in amyloid-positive individuals over age 60, and this offset between pathological change and cognitive change will be even greater for trials that begin in midlife. An opportunity may exist to shorten this delay from years to months by leveraging repeated testing, because amyloid-positive individuals may fail to benefit from these practice effects. Prevention trials may also benefit from a framework that conceptualizes amyloid PET status as a range rather than as dichotomous (i.e., positive or negative); amyloid PET levels in the subthreshold range (i.e., amyloid-negative) have been shown to correlate with future amyloid accumulation, and subtle differences in subthreshold amyloid among young individuals have been associated with recall memory performance. Tau PET holds similar promise as a biomarker that can be interrogated beyond positive or negative status, and focal signals in the medial temporal lobe and entorhinal cortex emerge as early as age 30-60 years. Other imaging biomarkers that could be relevant for midlife interventions include white matter hyperintensities, which demonstrate a more continuous accumulation with age than amyloid or tau, and blood–brain barrier (BBB) dysfunction, which has been demonstrated in disease context and is independent from amyloid pathology. In addition to imaging tools, proteomic approaches can be leveraged to investigate pathways that complement imaging-based biomarkers and research.

A national registry that recruits participants at early ages and leverages non-invasive data collection is a powerful opportunity to aggregate data, collect and store biospecimens (e.g., plasma), and identify potential participants for current and future data science projects, assay development, and intervention studies in midlife. Given that plasma biomarkers are inexpensive and easy to collect for the study of multiple biological pathways but lack spatial information and have unknown utility for early detection of disease, imaging biomarkers would be well-suited to validate plasma biomarkers in such a registry.

Blood-based Biomarkers: Opportunities for Early Prevention Trials of AD/ADRD

Adam M. Brickman, PhD, Columbia University

Biomarkers can be used for many purposes, including to detect or stage a disease, predict response to treatment, determine treatment efficacy, monitor treatment compliance, or monitor disease progression or recurrence. Although neuroimaging and cerebrospinal fluid

(CSF)-based biomarkers have been valuable tools in AD/ADRD research for several years, these biomarkers are limited by the low availability of radiopharmaceuticals needed for PET, high cost, reluctance of research participants to consent to procedures, and various contraindications, all of which leads to selection bias in research studies and a reduced ability to deploy these biomarkers in large trials. In contrast, blood-based biomarkers, which have primarily been used for diagnostic purposes, may be more easily deployed in large prevention trials because they do not require specialized equipment for collection, are less expensive than imaging- or CSF-based biomarkers, and can be collected as part of a routine procedure (i.e., blood draw). Moreover, blood-based biomarkers may be sensitive to pathophysiological effects before the onset of clinical symptoms and useful to test mechanistic hypotheses.

Blood-based biomarkers offer many opportunities for midlife early prevention trials. Plasma-based biomarkers may have prognostic utility over relatively short time periods to identify participants who are at high risk for decline or show early evidence of an intervention target for the purposes of trial inclusion. However, more prognostic data are needed in midlife and from diverse, community-dwelling adults before these biomarkers can be utilized effectively in this capacity. Blood-based biomarkers can readily be used, potentially, to detect target engagement, independent of whether the tested intervention is efficacious. Blood-based biomarkers may also be used as surrogate endpoints for testing intervention efficacy, but it is important to confirm that changes in the biomarker manifest as clinical benefit. Although most blood-based biomarkers lack sufficient evidence that they are more sensitive surrogate endpoints than cognition (especially in midlife) or that they are predictive of meaningful clinical outcomes, it is nonetheless valuable to collect biospecimens in parallel with clinical outcomes to continue study of these biomarkers. Large consortia (e.g., Biomarkers for Vascular Contributions to Cognitive Impairment and Dementia [MarkVCID]) and untargeted omics analyses are valuable sources of potential fluid biomarkers, including those that may be especially relevant for midlife studies (e.g., biomarkers related to vascular cognitive impairment). Emerging evidence suggests that plasma-based AD biomarkers—particularly phosphorylated tau (pTau)—are reliable across racial and ethnic groups and useful for detecting relevant information in midlife; however much more data from midlife and population-representative cohorts are needed to evaluate the validity and reliability of these biomarkers. Ultimately, scientific evidence—including the putative mechanism, fit for intended purpose, and prognostic utility—should drive the selection of biomarkers for a prevention trial.

Measuring Subtle Change in Psychological Processes: Challenges and Opportunities

Martin Sliwinski, PhD, The Pennsylvania State University

Measuring change is fundamental to establishing both target engagement and target validity. Changes in targets and endpoints (including surrogate endpoints) may not occur on the same timescale, and this disparity should be considered when designing assessment protocols. Much of the longitudinal evidence that supports the existence of modifiable risk factors for AD/ADRD entails relationships between static measurements of variables and outcomes from single time points; however, little evidence exists to support that *changes* in putative targets are associated with changes in cognitive function. The measurement of cognitive change is challenging

because this change is exceptionally subtle (less than 0.1 standard deviation units during a typical grant period), and this small effect size may place upper bounds on potential intervention effects that will be more stringent in midlife compared to studies in later life. Moreover, measurement error and confounds can overwhelm signals of cognitive change.

Change is an ongoing process that unfolds over multiple timescales. Although long-term changes are often of interest to clinical studies, long-term change occurs against the background of short-term change and variability (e.g., social activity, fatigue, work demands). One major challenge for measuring subtle change is temporal sampling error, which occurs when the influence of short-term variability is ignored, and both reduces measurement precision and obscures true long-term change. For example, working memory performance on a high- versus low-stress day can differ to a degree similar to the effects of 3-5 years of aging. Thus, fast-changing and difficult-to-control variables may overwhelm disease-related change or the effects of an intervention. Retest effects represent another major challenge for measuring subtle change. Retest effects are large and durable and can mask changes related to disease progression or interventions. For example, repetitive negative thinking (RNT) is a promising psychological intervention target for cognitive decline, yet recent evidence of cognitive change among older adults that scored high or low on measures of RNT at baseline only demonstrated differential change after retest effects dissipated (about 2-3 years after baseline). Higher-frequency, intensive measurement designs (e.g., measurement bursts) can improve the measurement of subtle cognitive changes. One such approach entails embedding cognitive tests into ecological momentary assessments (EMAs), which can boost reliability by reducing the influence of short-term variability. These intensive measurement designs allow for the assessment of variability over multiple timescales, can enable retest effects to be distinguished from actual cognitive change, and create an opportunity to examine digital biomarkers. More research on longitudinal target validation is needed to appropriately model change, which will require much more frequent measurement; this work would benefit from intensive measurement or micro-longitudinal designs and the utilization of pervasive technology and passive data streams to reduce participant burden.

Panel Discussion

Detecting and Defining Meaningful Change

There are many challenges to the detection of meaningful change in clinical prevention trials. A key consideration to address these challenges is the timeframe in which a trial will attempt to detect this change. Emerging evidence suggests that a well-designed study may be able to detect meaningful cognitive decline over the course of at least 1 year, even in midlife. It is also important that studies not focus strictly on measuring monotonic decline when the trait (e.g., cognition) is considered modifiable because study participants may voluntarily engage in behaviors that improve their functional performance. Another critical component to detecting meaningful change is the definition of meaningful change. If clinically meaningful cognitive decline (i.e., frank dementia) is the goal, for example, detection of this outcome is a tremendous challenge for midlife trials. However, functionally significant cognitive change—as opposed to *clinically* significant cognitive change—has been measured in individuals in their 20s

(e.g., in the Virginia Cognitive Aging Project). As prevention trials begin to start at earlier ages, it will also become increasingly valuable to identify meaningful functional outcomes that manifest earlier than the ultimate positive clinical outcome (i.e., the prevention of cognitive decline or AD/ADRD), because this clinical outcome may take years or decades of adherence to a lifestyle intervention to manifest; a more proximal functional outcome will promote adherence by reducing the delay between intervention initiation and observed benefit.

Developing and Utilizing Quantitative Biomarkers for AD/ADRD

Clinical studies represent an opportunity to both develop and utilize quantitative biomarkers for AD/ADRD prevention trials. The collection of biospecimens (e.g., blood) during a clinical study could provide large amounts of data for omics analyses that identify biomarkers (or combinations of biomarkers reflecting one or more biological pathways) that change with an intervention, even if those changes were not hypothesized a priori. The storage of these biospecimens would further allow the reuse of the same samples for biomarker development using assays that have yet to be developed at the time of sample collection—an important benefit for a rapidly evolving field. Blood samples are an especially appealing avenue for biomarker development given their minimally invasive nature.

Clinical studies can also leverage existing biomarkers to support their aims, although these biomarkers must be properly implemented to produce useful information. For example, study participants should be representative of the broader community to ensure that systematic biases are not impacting the interpretation of biomarker data. More research is needed to determine whether differences in race, ethnicity, or other demographic factors are reflected in biomarkers. Similarly, more work is needed to quantify the effects of age in fluid biomarkers so that typical biological aging processes can be distinguished from pathological decline. Age differences may also be reflected in digital biomarkers, which are important to recognize as studies increasingly rely on digital tools (e.g., smartphones).

Panel 2: Measuring Outcomes

Chair: Richard Gershon, PhD, Northwestern University

Standard Neuropsychology as an Outcome: Challenges and Opportunities

Yakeel Quiroz, PhD, Harvard University

Standard neuropsychological measures may not be optimal tools for AD/ADRD prevention trials, especially as these trials move into earlier asymptomatic stages. These measures were designed to identify or diagnose individuals with frank cognitive impairment, but participants in midlife prevention trials are generally included on the basis of biomarker-defined risk for AD/ADRD and would be considered cognitively normal on a standard neuropsychological assessment. A need therefore exists to improve neuropsychological measures so that they can better detect individuals at the greatest risk for AD/ADRD as well as track cognitive change—including intervention effects—over time. Opportunities to improve these measures include modifications that enhance the sensitivity of current standardized tests, application of

advanced psychometric methods to analyze currently available measures, and the development of novel measures.

Cognitive composites represent one approach to modify standard neuropsychological measures that have demonstrated an ability to track the earliest cognitive changes that are associated with underlying AD pathology. The Preclinical Alzheimer's Cognitive Composite (PACC), for example, is the most commonly utilized cognitive composite in preclinical trials and has been shown to detect cognitive decline in clinically normal individuals who are amyloid-positive versus amyloid-negative. Emerging evidence suggests that the inclusion of more cognitive domains beyond composite scores of episodic memory may provide greater sensitivity to detect β -amyloid ($A\beta$)-related cognitive decline; the POINTER modified Neuropsychological Test Battery (PmNTB) and the Alzheimer's Prevention Initiative (API) autosomal dominant AD (ADAD) Composite Cognitive Test Score are two examples of cognitive composites that assess cognitive function across multiple domains. Another opportunity to improve prevention trials is to develop new measures for the detection of subtle cognitive decline in unimpaired individuals at increased risk for AD/ADRD; the Memory Capacity Test, Face Name Associative Memory Exam (FNAME), and Short-Term Memory Binding test have all shown sensitivity to subtle cognitive changes that are associated with biomarker evidence of AD. The screening tools used to select individuals at increased AD risk can also be improved with these new measures. For example, home-based or unsupervised testing with new cognitive measures (e.g., FNAME) may differentiate amyloid-positive and amyloid-negative individuals. In addition, there is a critical need to develop and implement culturally and linguistically valid measures to increase diverse participation in clinical prevention trials, especially in light of evidence that these underrepresented populations are at greater risk to develop AD/ADRD.

Remote Cognitive Measurement as an Outcome

Richard Gershon, PhD, Northwestern University

The prevalence of dementia increases with age, but some cognitive decline is expected as part of normal aging beginning in early adulthood. To prevent and treat AD/ADRD-associated cognitive decline, it is essential to differentiate pathological cognitive changes from typical aging processes. However, this differentiation is hampered by a lack of sensitive assessment tools that can be easily and widely deployed in diverse research designs and populations. Currently available dementia screening tools and neuropsychological tests can detect varying degrees of cognitive impairment, but more sensitive assessments must be developed to detect earlier and subtler cognitive changes. To improve detection of preclinical cognitive change, assessments should also be affordable enough to enable frequent testing of cognition relative to an individual's own baseline. Furthermore, inexpensive surveillance tools would facilitate monitoring of potential interventions.

The Mobile Toolbox (MTB) for Monitoring Cognitive Function was designed to provide a smartphone-based library of mobile assessments of cognition as well as to enable delivery of additional variable measures (e.g., mood) that are known to be important covariates of cognition, including measures from the Patient Reported Outcomes Measurement Information

System (PROMIS). The MTB will feature a library of measures (all of which are available in Spanish) from multiple sources, including mobile formatted versions of NIH Toolbox for Assessment of Neurological and Behavioral Function (NIHTB) assessments, which have been widely validated in aging research and detection of cognitive change from ages 3 to 85. The included measures have demonstrated evidence of sensitivity to age and/or age-related neuropathology. Furthermore, the tests focus on stimuli that can be randomly generated or drawn from a large item pool to avoid stimuli-specific practice effects and are adaptable to mobile testing (with an emphasis on EMA or burst designs). The MTB will be open source and will support researchers to contribute or select assessments to the library; manage study and assessment protocols; and receive, store, aggregate, analyze, and responsibly share data collected on smartphones. Researchers can also create, deploy, monitor, and analyze complete smartphone-based studies using existing tools within the MTB ecosystem. Smartphone platforms (in particular iOS and Android) offer many advantages for mobile monitoring of cognitive change, including increasingly prevalent use relative to regular computers (including in underrepresented communities in research), an effectively free distribution model for apps, and many available sensor types. The constructs employed by the MTB are currently being validated against the NIHTB and external gold standards in a general population sample of adults aged 20-85. Public dissemination of the MTB is anticipated in 2023.

Found Measures of Cognition

Lauren Nicholas, PhD, University of Colorado

Dementia entails a loss of cognitive functioning that interferes with a person's daily life and activities. This interference with daily life produces a symptom trail—including financial (e.g., forgetting to pay bills, impetuous purchasing), behavioral (e.g., personality changes, Google searches for symptoms), and mobility (e.g., car accidents, new reliance on rideshares, changes in actigraphy data) consequences—that could provide insight into the earliest stages of AD/ADRD. Patterns of these consequences may already be measured in large data sources that are not traditionally used in research. For example, the linkage of Medicare claims data with Equifax/Federal Reserve Bank of New York Consumer Credit Panel data has enabled the study of whether AD/ADRD has a unique financial presentation. This linkage enables exploration of changes in rates of adverse financial events as a function of time from AD/ADRD diagnosis (or lack of diagnosis) and may reveal whether financial events at certain points in time may be predictive of future AD/ADRD. Emerging evidence from this analysis has demonstrated an increased risk of credit account payment delinquency and new subprime credit score approximately 6 years and 2.5 years, respectively, prior to AD/ADRD diagnosis. These risks were not observed for other health conditions (e.g., glaucoma, hip fracture) but are not necessarily exclusive to AD/ADRD. Individuals with high risk scores generated from single timepoint credit data are more likely to develop AD/ADRD than those with lower risk scores, although more data are needed to refine these scores for diagnostic purposes.

Found measures such as financial data have many strengths for AD/ADRD research: they (1) may already be present in existing datasets, (2) are relatively low cost and low burden relative to other methods of collecting cognition measures, (3) can be very high frequency and sensitive

to narrow time windows, (4) have near universal participation useful for generating large samples with high retention rates, (5) can improve insight into how dementia impacts daily living, and (6) are useful for monitoring symptoms and devising tools for AD/ADRD detection and protection. However, the use of found measures is also accompanied by challenges. Privacy concerns exist with found measures, because they are often sensitive and therefore involve lengthy foundational work (e.g., obtaining permissions) to link additional sensitive data for obtaining AD/ADRD status that may be difficult to complete during typical research funding timelines. In addition, found measures often require linking information from multiple disparate data sources that may not be interoperable, which requires a wide range of skill sets and institutional knowledge.

The design of prospective studies with found measures is challenging because these prospectively collected measures may become obsolete. In addition, automation of activities captured by found measures (e.g., automatic bill payments, self-driving cars) may mask the effects of cognitive decline. More validation studies are needed to identify the most promising metrics, the benefits of which will likely extend to many other diseases and interventions that seek to improve functional status. This foundational research could be done in parallel with clinical trials. An opportunity to expand and improve the use of found measures in research is to obtain prospective consent for a Health and Retirement Study (HRS)-type dataset with passively collected functional measures of daily living. Another opportunity may be to partner with marketers and technology companies that are likely already performing this sort of passive data collection. Supporting the development of a new dataset that can be used by many research teams to study AD/ADRD, in addition to other diseases, could yield many new discoveries similar to the success of HRS.

Panel Discussion

Developing New Measures

Emerging evidence suggests that found measures of cognition could be a valuable resource for the prediction of AD/ADRD outcomes; however, these measures are not yet sufficiently developed for use as diagnostic tools. An opportunity to improve the sensitivity and specificity of found measures as well as other outcome measures is to combine multiple measurement domains (e.g., visuospatial and psychomotor) to enhance predictive abilities for research. These different measurements could capture non-cognitive change that is predictive of AD/ADRD and would be differentially sensitive to common research challenges such as practice effects because these effects are domain-specific. An ongoing challenge for the development of robust outcome measures is the lack of access to valuable data types. Informant data, for example, would improve predictive models but are not represented in traditional health data capture systems. Likewise, public-private partnerships could grant access to nontraditional data sources that are likely to already be conducting passive data collection that could inform AD/ADRD prediction. Although credit data may be easier to access for research because a market already exists in which those data are sold, other private-sector data sources may be more difficult to pursue and would require more investment of time and resources from private partners.

Detecting AD/ADRD Earlier in Life

To prevent cognitive decline and AD/ADRD, measures are needed that can detect change many years before diagnosis. One barrier to developing these measures at earlier ages is the recruitment of a sufficiently large sample. More data over longer follow-up periods from large data linkage efforts are needed to improve the power and precision of midlife prevention trials. To obtain these longitudinal data, longer-term funding mechanisms and/or access to greater volumes of retrospective data will be essential. Another challenge is the knowledge gap related to when the earliest and subtlest cognitive decline occurs relative to AD/ADRD diagnosis. One opportunity to address this knowledge gap is to test outcome measures in cohorts with known high risk for AD/ADRD, such as individuals with autosomal dominant AD. These cohorts may also be leveraged as pathological controls when measures for midlife assessments are developed and tested because these individuals are likely to be diagnosed earlier than people with other forms of AD/ADRD, which reduces the confounding influence of later life aging processes in older control groups.

Panel 3: Issues in Intervention Development: How to Design the Most Informative Trials

Chair: Linda M. Collins, PhD, New York University

Developing Effective, Affordable, Scalable, and Efficient Interventions

Linda M. Collins, PhD, New York University

The classical treatment package approach, in which interventions are evaluated via randomized controlled trials (RCTs), is hindering the progress of intervention development. Although RCTs are very useful for assessing the performance of entire intervention packages, they do not provide insight into individual intervention components. Because the classical treatment package approach establishes efficacy first without considering other factors, this approach often produces interventions that are not affordable, scalable, or efficient and can only be made so by ad hoc modification of intervention components. However, without insight into the individual components of an intervention, these modifications could unintentionally undermine effectiveness by eliminating critical components.

The MOST framework is an approach to intervention development that is rooted in optimization: the process of identifying a strategic balance of effectiveness against affordability, scalability, and efficiency (*EASE*). Optimization is not about developing the absolute best intervention, but rather the best intervention that is subject to realistic constraints (e.g., willingness to pay, available staff time, and tolerable level of complexity). The MOST framework has three phases: preparation, optimization, and evaluation. In the preparation phase, intervention models are conceptualized, candidate components are identified, and a specific definition of *EASE* is set as the optimization objective. In the optimization phase, the intervention is built by conducting one or more optimization trials that use designs other than the classical RCT (e.g., factorial experiment or sequential multiple assignment randomized trial [SMART]). The results of these optimization trials are used to identify the set of components required to meet the optimization objective. In the final phase, the optimized intervention can

be evaluated in an RCT. Two fundamental principles guide this process: the continual optimization principle, which states that interventions must continue to be improved over time, and the resource management principle, which states that the investigator must select the most efficient experimental design for the optimization trial. The MOST framework can accelerate progress toward prevention of AD/ADRD by providing more insight into why interventions are successful and illuminating logical next steps to continually improve interventions. For example, where the classical treatment package approach might identify one effective intervention, the MOST approach could define the specific components of that intervention that contribute to its success (as well as successful components of intervention packages that might have failed along the way and could be salvaged) and suggest opportunities to make the intervention more affordable, scalable, or efficient without undermining effectiveness.

Who Do We Really End Up Recruiting into Our Trials and Research?

Melinda C. Power, ScD, George Washington University

Two major study types are usually leveraged for biomedical research: observational studies and clinical trials. Observational studies are often designed to recruit a sample that is representative of a defined population (e.g., community, clinic, or administrative unit) at baseline. In contrast, clinical trials typically recruit a convenience sample that meets a large set of inclusion and exclusion criteria that reflect ethical, safety, and feasibility concerns. Relatively few participants in observational studies meet the inclusion and exclusion criteria that determine eligibility for clinical trials. Examples of this challenge include the Alzheimer's Disease Neuroimaging Initiative (ADNI), a highly selective sample designed to emulate a clinical trial population, and the Atherosclerosis Risk in Communities Study (ARIC), a community-based sample designed to be representative at baseline. Sample characteristics differed substantially across the two studies, illustrating that who is included in the sample differs. Furthermore, a comparison of exposure-outcome associations demonstrated that about one-third of all considered associations differed statistically and about one-half of all considered associations differed qualitatively across the studies. These differences indicate that study conclusions are often sample dependent. Thus, results from highly selective populations may not generalize to more representative populations and vice versa.

Research samples can lack sufficient representation of important subgroups to quantify subgroup effects. Historically, both observational studies and RCTs of aging and dementia recruited samples that were predominantly non-Hispanic white (although this is starting to change). However, the burden of disease is known to vary by race and ethnicity, with greater burden in minoritized groups. To identify the drivers of persistent racial disparities and make progress toward health equity, studies need sufficient representation of different subgroups.

How samples are recruited and data are collected also influences who is identified as having dementia. Dementia can be ascertained by various methods, including the criterion standard (i.e., in-person research-based assessment), algorithms based on cognitive testing and other data, and use of administrative data sources (e.g., death certificates, EHRs, or Medicare claims).

The sensitivity, specificity, and accuracy of non-criterion standard ascertainment methods often vary systematically based on sample characteristics. For example, a comparison of use of Medicare claims to criterion standard ascertainment demonstrated that use of Medicare claims to identify persons with dementia may miss almost one-half of persons living with dementia, and these persons are more likely to be younger, male, and Black; have lower educational attainment; and have diabetes. Similarly, a substantial portion of persons were incorrectly identified as living with dementia, and these individuals are more likely to be older and male, have higher educational attainment, and have diabetes. Thus, different ascertainment methods can yield different results, and study heterogeneity may reflect measurement artifacts rather than true findings. Although the use of algorithms or administrative data sources may enable study of more representative samples, it may also induce undesirable measurement error that can bias study findings.

Recruitment processes are not the only determinant of sample composition. Loss to follow-up is an issue in both observational studies and clinical trials. Loss to follow-up does not occur randomly, because persons lost to follow-up are often older, have more cognitive impairment and functional limitations, have more health conditions, and have more risk factors for poor health. As a result, effect estimates may not generalize to persons who are lost to follow-up. If loss to follow-up is related to exposure and outcome, it may also induce selection bias.

Sources of heterogeneity across populations must be considered to make observational data and clinical trial data more relevant to each other. The application of missing data procedures is one opportunity to reduce biases and close the gap between trial designs. Overall, researchers should think carefully about the design and analysis features of observational studies and trials to render the results of both more useful.

Large Trials in the Health Care System to Detect and Prevent Dementia

Jason Doctor, PhD, University of Southern California

The United States health care system is fragmented and is therefore a challenging venue for clinical trials. Each EHR system is different, and health care data can be difficult to analyze and utilize for long-term follow-up. While data that have financial or legal uses (e.g., diagnoses or orders) are generally well-documented, other data may be poorly structured in EHRs and patient-reported outcomes (PROs) are not universally used. The Patient-Centered Scalable National Network for Effectiveness Research (pSCANNER) Network is an effort to connect and standardize EHR data throughout the United States in order to better facilitate RCTs in health care systems.

Research in large health care systems has revealed multiple behavioral insights that could improve care. First, EHR data studies have demonstrated that physicians experience decision fatigue and make poorer prescribing choices later in the day after making repeated decisions, suggesting that breaks would be beneficial for physicians' decision-making abilities. EHR systems can also be used to recruit and track study participants if assessment mechanisms are incorporated into the EHR system. A study that leveraged EHR recruitment and tracking demonstrated that loss aversion-based incentive strategy can be effective for exercise

interventions; issuance of lottery tickets whose likelihood of success was linked to intervention adherence was a more effective incentive than a control condition that included both the lottery and a fixed payment. Another behavioral insight from studies in health care systems is the power of public commitments. In one trial, inappropriate prescribing behaviors decreased by almost 20 percent when physicians worked in exam rooms with a printed commitment not to overprescribe posted on the wall. Inappropriate prescribing has also been studied using factorial designs in health care systems. Evidence from these studies demonstrates that behavioral interventions that notify physicians who are among the top performing prescribers (i.e., peer comparison or social norms intervention) or that prompt physicians to justify their prescribing behavior (i.e., social justification or accountability intervention) are both effective methods to reduce inappropriate prescribing. In addition, interventions that make the harms of inappropriate prescribing more apparent—for example, by making non-judgmental information about opioid deaths in a physician’s practice more readily available—reduced inappropriate prescribing by almost 10 percent in one clinic-based study.

Although the amyloid and tau hypotheses of AD/ADRD have long been dominant, approximately 200 drug trials have failed in the last 30 years, and a causal understanding of AD/ADRD pathology has not emerged. Newer theories of AD/ADRD pathology create space for health care system interventions. These interventions may focus on patient behavior (e.g., exercise) or common clinical practices. For example, interventions could encourage more judicious prescription of common medications linked to dementia (e.g., anticholinergics and benzodiazepines) or medications that exacerbate the behavioral and psychological symptoms of dementia (e.g., antipsychotics). A health system improvement trial could also be utilized to improve detection of AD/ADRD (e.g., by reimbursing comprehensive assessments).

Panel Discussion

Promoting Iterative Optimization of Interventions

The MOST approach to intervention development is predicated on iterative discovery. However, barriers exist to supporting MOST trial designs with existing funding mechanisms. Although more than 100 MOST projects have been funded by various NIH funding mechanisms, the traditional 5-year grant period places challenging time constraints on an inherently iterative optimization framework. Moreover, it is difficult to propose a series of trials in which fundamental details about later trials would be defined by earlier trials using current grant mechanisms. For example, initial trials may generate data that suggest that either an optimization trial or an RCT is an appropriate next step, but this next step must generally be defined in a grant proposal before the initial trial is conducted. Reviewers also tend to place more value on RCTs than optimization trials, which may disadvantage proposals for optimization trials despite the substantial contributions that optimization trials make to the field. Reviewer perspectives on the value of iterative optimization trials and the structure of currently available funding mechanisms likely need to evolve to promote intervention development via the MOST approach. Such an evolution will hasten scientific progress toward development of non-pharmacological approaches to the early prevention of AD/ADRD.

Composition of Prevention Trial Cohorts

Some differences between trial populations and the general population are not inherently problematic; individuals that qualify for a clinical trial often must have the risk factors necessary to meet inclusion criteria while being free from the contraindications that violate exclusion criteria. However, unmeasured differences between trial populations and the general population may drive some aspects of trial results. Furthermore, clinical trials must continue to strive for more representative samples. Modern clinical trials have better racial and ethnic diversity than past clinical trials, although continued effort is needed to make trial populations more representative and improve the generalizability of clinical trial results. In addition to recruiting more representative samples, clinical trials could leverage transportability algorithms to project how results may translate to other populations. Importantly, transporting estimates requires sufficient sampling in order to provide appropriate subgroup estimates; without these estimates, results may not be accurately estimated in a different population that has more or less of a given subgroup. Another challenge for clinical trial recruitment is the potential for bias toward recruiting individuals who are especially motivated to participate in the trial, which can create a disconnect between intervention efficacy and effectiveness when the broader population is unwilling to adhere to the intervention as faithfully as the trial population. Furthermore, these participants may differ on other important (if less obvious) characteristics. Therefore, clinical trials should be pragmatic and consider whether interventions can be feasibly implemented at scale rather than solely focused on efficacy.

Incentives for Midlife Prevention Trials in Large Health Care Systems

A key advantage of performing midlife prevention trials in large health care systems is the availability of EHRs and the capacity to positively influence physician behavior (e.g., by encouraging judicious prescribing practices, especially for benzodiazepines, antipsychotics, and anticholinergics). However, incentive structures are needed to encourage large health care systems to assume the cost of prevention trials; one such opportunity would be to increase the prevalence of Medicare Advantage programs.

Panel 4: Non-intervention Development: Generating Causal Evidence from Observational and Non-clinical Studies to Inform the Prevention Research Agenda

Chair: Maria Glymour, ScD, MS, University of California, San Francisco

Causal Inference from Observational Data

Sonja Swanson, ScD, Erasmus MC

Early midlife prevention trials for ADRD are of substantial clinical and public health interest but are hindered by inherent challenges to feasibility and timeliness because outcomes require years or decades to manifest. An alternative approach to generating causal evidence for ADRD prevention is to carefully design and conduct an observational study that emulates the ideal randomized experiment. This alternative can also complement existing trials: if the ideal trial is infeasible, results from observational data may be benchmarked against a feasible related trial.

A critical first step toward emulating a target trial is to carefully articulate the protocol elements of the target trial. This process of specifying the target trial's elements is essential and requires epidemiologic, methodologic, substantive, and data-specific expertise to understand the feasibility of emulating such a trial with the available data. In addition, this process must include end users and stakeholders to ensure that the emulated trial is itself informative for meaningful questions. Given a particular data source, the target trial may still be a compromise, although recognition of this possibility helps to contextualize the totality of information that is gained. Observational studies should be conducted as closely as possible to the target trial with strategies in place to avoid, mitigate, understand, and quantify biases. Furthermore, observational analyses should be transparent in presentation and interpretation with respect to underlying assumptions.

Observational data often allow for more agility in recruitment of diverse populations. This advantage offers more opportunities to study heterogeneity across subgroups, identify the subgroups that would benefit most from an intervention, contextualize disparities, and address inequities. However, disparities in study participation will still exist in observational data, and measurement of relevant criteria in all available data types may be challenging. Another advantage of observational data is the ability to efficiently study multiple related or joint interventions, whereas it is often infeasible to substantially increase the number of treatment arms in a clinical trial. However, individuals following different interventions in an RCT are expected to be comparable at baseline. Observational data may be used to emulate this randomization by assuming random assignment within levels of measured covariates (i.e., adjusting for measured confounding) or by finding a natural source of variation. Nonetheless, randomization remains the fundamental difference between observational studies and clinical trials. Thus, researchers need to embrace falsification strategies and sensitivity analyses in order to assess the sensitivity of study conclusions to underlying assumptions. It is also important to align three events—treatment assignment, eligibility criteria application, and start of outcome recording—to the same “time zero” when emulating a trial. An unclear or misaligned time zero could create similar selection biases as would occur if participation was conditioned on post-randomization events in a trial.

Multiple challenges are shared by both observational studies and clinical trials, including the need to focus on patient-centered effects and the inevitability of non-adherence, loss to follow-up, and deaths as competing events in studies of sustained interventions. Recognition of these shared challenges informs best practices across both RCTs and observational studies and facilitates comparison of results, gaps, and opportunities.

Quasi-Experimental Methods

Emilie Courtin, PhD, MSc, London School of Hygiene & Tropical Medicine

Quasi-experimental approaches can generate robust causal evidence from observational data to inform the prevention research agenda. A quasi-experiment leverages exogenous and uncontrolled sources of variation in intervention exposures and can be designed in multiple ways. The instrumental variable approach leverages a source of variation that is unrelated to

other determinants of the outcome of interest; for example, distance to school may be used as an instrument to estimate educational attainment in order to examine the impact of educational attainment (as predicted by the instrument) on health outcomes. The regression discontinuity design approach leverages an arbitrary discontinuity in the probability of being treated with an intervention; for example, probability of exposure to a schooling reform may be arbitrarily impacted by birth year. The difference-in-differences approach compares outcome differences between a group exposed to an intervention and a control group; for example, this approach may compare people in states that did or did not enact policies that improved funding of schools. Hallmarks of a strong quasi-experiment include the presence of an exogenous intervention and the intention that the intervention should affect an outcome. There are many opportunities to leverage quasi-experimental designs to build the evidence base for AD/ADRD prevention. For example, quasi-experiments in this field could investigate gene-by-environment interactions between *APOE* genotype and educational opportunities, examine the influence of clinical decisions to initiate blood pressure screening based on defined thresholds (e.g., medical cutoffs or health policy differences) in midlife, explore geographical variation in environmental policies (e.g., noise and pollution), and utilize spousal death as an instrument to study the impact of depressive symptoms on later life cognitive function. Throughout the lifecourse, quasi-experiments may be used to investigate causal questions about the determinants of health inequalities in cognition and assist with the identification of effective interventions.

Quasi-experiments may generate causal evidence when RCTs are not possible (e.g., due to ethical constraints or to evaluate policies and interventions that occurred in the past). A robustly designed quasi-experimental approach has high internal and external validity and can mimic randomization in a real-world setting. Therefore, the findings of quasi-experimental studies are theoretically more generalizable than RCT results and can evaluate the effectiveness of an intervention rather than solely assess efficacy. In addition, quasi-experimental approaches may be more efficient than RCTs (depending on data availability) for midlife AD/ADRD prevention studies because they are relatively quick to conduct with low associated cost and long follow-up periods. However, quasi-experimental approaches also come with challenges, including a higher risk of bias than RCTs. These approaches are also inherently rooted in discovery rather than prospective planning because the interventions and policies evaluated by quasi-experimental methods must have occurred in the past. In addition, quasi-experimental studies rely on available data and require large sample sizes to distinguish signal from noise in routinely collected datasets. Interventions may also not be well-defined in quasi-experimental studies because eligibility and implementation fidelity may not always be clear. Importantly, the Target Trial Framework has recently been adapted for quasi-experiments to provide a systematic basis to assess the plausibility of the assumptions and claims that must be made in these approaches. A multidisciplinary approach is key to the success of using quasi-experimental methods for prevention research because these quickly evolving methods have been developed by social scientists and training is needed to equip public health researchers with the necessary tools. Transparency and replicability can further strengthen trust in these methods. Moreover, establishing registration, falsification, and triangulation as norms in the prevention research field would support the use of these methods.

Measuring Effects of Treatment Variation in Healthcare Systems and Their Effects on Cognition

Rachel Whitmer, PhD, University of California, Davis

Several critical life stages have been identified for which some exposures may have a greater impact on brain health. Midlife has emerged as a particularly important life stage for the AD/ADRD prevention research agenda in part because neuropathology can be observed decades before symptom onset. Thus, the development of prevention strategies for AD/ADRD require a lifecourse approach to assessing risk.

Studies that leverage health care data are capable of powerful observational insights for AD/ADRD prevention. Health care data not only encompass various data types (e.g., direct clinical measurements, surveys, treatment histories, and comorbidities) but also can be rich sources of longitudinal data when members have been part of a health care system for many years. These longitudinal data collected at different times of peoples' lives allow researchers to ask innovative questions about risk and resilience for AD/ADRD. For example, a study of long-term members of Kaiser Permanente in California produced the first longitudinal analysis that associated heavy smoking in midlife with long-term risk of all-cause dementia. Similar approaches have been utilized to demonstrate an association between early adulthood hypertension status in females and greater risk for all-cause dementia. A major advantage of health care system data is the ability to examine how the timing of risk factor exposure influences outcomes. For example, these data can reveal when cardiovascular risk factors emerge in different groups and how the timing of that emergence is associated with late life cognition. Health care data systems will also generally include geographical information that can be linked to other data sources to investigate the interaction between environmental (e.g., air pollution or food density) or psychosocial (e.g., socioeconomic status) exposures and health over time.

AD/ADRD prevention trials may incorporate health care data to inform the timing of interventions and validate assessment tools. For example, midlife health care system data were used to demonstrate that a combination of modifiable risk factors at midlife is highly predictive of dementia more than three decades later, as well as to validate and improve the Cardiovascular Risk Factors, Aging, and Incidence of Dementia (CAIDE) risk score. Moreover, these data can be used to enrich trial populations based on risk and resilience insights obtained in midlife (e.g., CAIDE risk score) and to assess how risk factors should be interpreted at different stages of life. When incorporating health care system data, it is important to recognize that equal access to a health care system does not imply equal usage of a health care system, as barriers to care will still exist for some member groups. Furthermore, standards of care and screening trends change over time, which may impact prevention trials that utilize these data.

Chronic Conditions and Risk of AD/ADRD: Advancing Prevention Analyses to Inform Population Health

Julie Zissimopoulos, PhD, University of Southern California

In the absence of treatments to prevent, treat, or cure AD/ADRD, pharmacological therapies that are currently used to treat high prevalence non-AD/ADRD conditions may be able to reduce the population burden of AD/ADRD as well as racial or ethnic disparities in risk for dementia. Because conditions such as hypertension are associated with AD/ADRD risk, it is feasible that the drugs used to treat these conditions (e.g., anti-hypertensives [AHTs]) may act on dementia via control of the associated condition (e.g., hypertension) or that these drugs may have pleiotropic effects that act directly on AD/ADRD pathology. Thus, potential approaches to reducing AD/ADRD risk include repurposing drugs, adjusting treatment across drug classes, initiating treatment or achieving better adherence to treatments, and reducing the use of drugs associated with increased AD/ADRD risk. Real-world evidence (RWE) can provide valuable support for regulatory decision making—including the approval of new indications for approved drugs. However, its use requires a multidisciplinary research group, a thorough understanding of data quality, appropriate study design, rigorous methods to mitigate bias, and appropriate expectations about what RWE can and cannot explain.

Claims data are a source of RWE with many strengths, including large sample sizes that allow examination of specific drug classes and the heterogeneity of effects across populations. Challenges to the utilization and interpretation of claims data include various threats to validity, such as reverse causation, confounding, selection bias, misclassification, and issues with rigor or robustness. In addition, the availability of a lot of data can lead to data mining rather than hypothesis testing, although collaboration with clinical and basic scientists supports hypothesis-driven research with claims data. For example, one collaboration generated a hypothesis that renin-angiotensin system (RAS)-acting AHTs may confer more protection against AD/ADRD than other AHTs. Moreover, the researchers hypothesized that sex and race differences would be observed on the basis of estrogen activity on RAS and variable endogenous levels of sodium and renin. Common sources of biases in claims data were mitigated by (1) requiring at least two diagnoses of AD over time to avoid misclassification of outcome, (2) defining AHT users based on days of drug supply received over two consecutive years, (3) assuming that any misclassification bias is the same across drug classes, (4) requiring no prior diagnosis of AD/ADRD or mild cognitive impairment to avoid reverse causation, (5) controlling for demographics to address confounding, and (6) restricting the sample to beneficiaries who survived and were continuously enrolled in fee-for-service (FFS) Medicare through the end of the study period to address selective attrition. A key assumption of the study was that sorting into type of AHT was uncorrelated with AD risk after adjusting for confounders (i.e., no unobserved confounders). The researchers found strong evidence consistent with the hypothesis that angiotensin receptor blockers (ARBs), but not ACE inhibitors (ACEIs), reduced risk of AD compared to non-RAS-acting AHTs. In addition, sex differences consistent with the hypothesis of estrogen-RAS interaction were observed, although no strong evidence for differential effects across racial groups was found.

Longitudinal, nationally representative RWE data (e.g., HRS) can be leveraged to estimate demographic, health, and socioeconomic factors associated with dementia onset over time. Dynamic microsimulation can then be employed to project life expectancy and dementia prevalence under difference scenarios of improved population health (e.g., reducing diabetes or cardiovascular disease risk) and compared to a hypothetical new AD/ADRD treatment. A key assumption of this approach is that the measured associations in models that inform the dynamic microsimulation will remain valid in the future. Evidence from a dynamic microsimulation suggests that reducing incidence of hypertension or diabetes, as well as new AD/ADRD treatments, would all increase life expectancy. However, only new AD/ADRD treatments were estimated to reduce the incidence of dementia at ages 85-86 relative to the status quo in this analysis. Similarly, only an intervention that acts directly on dementia was estimated to reduce population levels of AD/ADRD 20 years in the future. This kind of modeling can be used to gather and interpret data from clinical trials to project health impacts and quantify the costs and value of treatment for individuals, care partners, and the overall population.

Panel Discussion

Analysis of Subgroup Differences

Although disparities in AD/ADRD risk are widely acknowledged, analysis of subgroup (e.g., race, ethnicity, sex, socioeconomic status) differences remains a challenge. Approaches such as quasi-experiments can be used to interrogate the effects of a policy change on dementia risk; however, these analyses do not often include a systematic investigation of how those policies may reflect societal inequity (e.g., structural racism) that has consequences for long-term health. Examination of subgroup differences is hampered by increasingly limited power as more subgroups are added to an analysis, which requires very large sample sizes to address. Similar challenges exist for the analysis of different causes of AD/ADRD. The high prevalence of mixed pathology has resulted in much of the prevention research agenda focusing on all-cause dementia rather than distinct subtypes. Moreover, postmortem neuropathology does not perfectly correlate with clinical presentation, further complicating the ability to study or correctly identify subtypes of AD/ADRD. While very large samples are necessary to obtain subgroup-specific results with robust statistical power, the quality of phenotyping often decreases when larger samples are obtained. Targeted studies that aim to gather high-quality measurements from a subsample, as in the HRS Aging, Demographics and Memory Study (ADAMS), are one example of a study design that may provide more insight into specific biological mechanisms underlying changes in cognition. European countries that are currently working to augment administrative datasets (e.g., United Kingdom or Denmark) may also serve as exemplars for data collection in large samples that enable quasi-experimental research with high power.

Moving Beyond Traditional RCTs

Methodologies besides traditional RCTs can both deliver rigorous evidence when RCTs are not feasible and enhance the design and conduct of future RCTs that will be more effective and deliver evidence more quickly. To adopt these tools, training for some methods that are not

part of the traditional canon of epidemiologic research must be supported. For example, the inclusion of the fundamental principles of emulating a target trial could be made a cornerstone of future epidemiological research education; efforts to promote the creation, recognition, and utilization of quasi-experimental methods would also deliver a new and powerful set of tools to AD/ADRD research when formal, researcher-initiated RCTs are not feasible. Furthermore, the guiding principles of these nontraditional methods could inform one another to amplify potential knowledge gains. For example, reanalysis of existing observational studies could theoretically be used to assess the contributions of individual components of an intervention (similar to optimization trials in the MOST framework); however, this approach would require full transparency and a thorough understanding of how biases may propagate as decisions are made in iterative optimization trials, as well as a larger volume of observational data that can speak to individual components of interventions.

A greater emphasis on multidisciplinary research will be essential to proper adoption of nontraditional approaches as more diverse expertise, methods, and data types become critical components of new trial designs. Research that operates successfully at the intersection of various fields (e.g., clinical and basic neuroscience, economics, epidemiology) may involve a steep learning curve for all involved collaborators, but the intellectual challenge will enable insights that would not be possible with siloed efforts. As multidisciplinary studies become more commonplace, it will also become more essential to train researchers in effective communication of science to different audiences in order for their work to be received well by reviewers, who may not have comprehensive backgrounds in the disciplines that inform any given study.

General Discussion

The panelists broadly agreed that both more data and better methods to leverage existing data are needed to support AD/ADRD prevention trials in midlife. Panelists therefore highlighted the following gaps and opportunities in the prevention research agenda for AD/ADRD that could address these needs.

Larger, More Representative Samples

To successfully develop and test interventions that may prevent AD/ADRD or cognitive decline, more comprehensive data are needed from larger cohorts over longer time periods. Moreover, these cohorts must be more representative of the target population in order to sufficiently assess subgroup differences that are widely accepted to exist but are poorly understood. Researchers must consider approaches to remove barriers to recruitment as well as retention in order to build a solid evidence base for determining who is at the highest risk for AD/ADRD and evaluating potential interventions. Importantly, these barriers may be different for different study types (e.g., observational studies, RCTs, health care system–based studies) because current cohorts differ systematically across these study types, which may impact study results.

Multidisciplinary Research Approaches and Modern Causal Inference Methods

Different expertise is needed to effectively recruit target populations; develop high-frequency collection tools; validate new biomarkers; create, recognize, and analyze quasi-experiments;

and design observational analyses to identify causal effects. However, *all* of these tasks are necessary to design effective prevention trials for AD/ADRD. Thus, a multidisciplinary approach is essential to the acquisition of more comprehensive data for the prevention research agenda. A consortium with a breadth of expertise would facilitate large-scale data collection and analysis as well as trial design innovation (i.e., study designs other than traditional RCTs, including emulated trial designs and quasi-experiments), which will help researchers to develop interventions more rapidly by supporting decision making based on currently available data as well as highlighting what data are missing and needed.

Trials Within Large Health Care Systems and Similar Settings with Long Term Data

Large health care systems represent a valuable opportunity to collect longitudinal data from large populations. Furthermore, the collection of these data in a health care system is more cost-effective than a full RCT with long follow-up periods. Health care systems are also important settings for testing physician behavior interventions as well as for gathering data on exposures that are too dangerous to assess in a clinical trial setting; for example, health care system data could be leveraged to investigate the influence of severe low blood sugar episodes on dementia risk.

Infrastructure to Support Biomarker and Tool Development

The development of biomarkers and tools for AD/ADRD prevention research is complicated by the long delay between biomarker collection and outcome; because technologies evolve rapidly, the biomarkers and assessments utilized at the start of a lengthy trial may not be as useful years later when outcomes are assessed. A registry that collects cognitive assessments as well as biospecimens would support tool development by providing a venue for long-term data storage with samples that can be retested as new biomarker assays are validated. For example, ADNI has performed multiple iterations of analysis on banked CSF samples so that it could take advantage of new assays that were not available at the time of sample collection. One important goal for biomarker and tool development is cost efficiency; the development of useful neurocognitive assessments or fluid biomarkers could greatly reduce the costs associated with large-scale trials compared to more expensive tools such as amyloid PET imaging. Another priority is to develop biomarkers that can provide more rapid insights into whether an intervention is working. These biomarkers would be especially useful for MOST trial designs, which rely upon a research team's ability to quickly evaluate and act upon information about intervention component successes and failures for subsequent optimization trial iterations.

Appendix A: Meeting Agenda

Workshop Chairs: Jonathan W. King, PhD and Maria Glymour, ScD, MS

Steering Committee: Linda M. Collins, PhD; Jason Doctor, PhD; Richard Gershon, PhD; and Beth Mormino, PhD

Non-pharmacological interventions to prevent age-related Alzheimer's Disease or Alzheimer's Disease Related Dementias (AD/ADRD) are of great interest to scientists, policy makers, and the public. Of particular interest would be prevention interventions beginning in midlife that could eliminate disease burden completely, as well as non-pharmacological interventions that would avoid the risks and costs of chronic drug regimens. Nevertheless, the "naïve" approach to the non-pharmacological prevention of AD/ADRD beginning in midlife, a decades long randomized clinical trial, would face immediate and significant design challenges. The National Institute on Aging's Division of Behavioral and Social Research is convening a two-day virtual workshop to address these core methodological challenges with respect to the populations we should be selecting, the intervention targets we are attempting to hit—including how best to measure target engagement—and the actual and specific outcomes we are seeking. In addition, we will examine the overall process of intervention development and address how best to accumulate evidence to support intervention strategies that will be effective in populations of interest.

Day 1: May 24, 2021

- 1:00 pm Welcome and Introductions – Lis Nielsen, Eliezer Masliah, and Maria Glymour
- 1:10 pm State of the Science and Precip for the Meeting – Jonathan W. King
- 1:25 pm Break

Panel 1: Assessing Target Engagement

Chair: Beth Mormino

This panel will offer an introduction to (the state of evidence for the malleability of) potentially modifiable molecular, neural, or psychological processes on the early trajectory to AD/ADRD that may serve as targets for preventive interventions for AD/ADRD in early life or midlife. Talks will focus on how well-suited our measures of these processes are for inclusion in prevention trials and what sort of evidence would be needed to ascertain that an intervention was able to engage these processes.

- 1:30 pm Introduction – Beth Mormino
- 1:35 pm Imaging Biomarkers That Could Index Change – Beth Mormino
- 1:55 pm Blood-based Biomarkers: Opportunities for Early Prevention Trials of AD/ADRD – Adam M. Brickman
- 2:15 pm Measuring Subtle Change in Psychological Processes: Challenges and Opportunities – Martin Sliwinski

2:35 pm Break

2:40 pm Discussion: What more do we need to know about targets and measures of target engagement?

Panel 2: Measuring Outcomes

Chair: Richard Gershon

This panel will consider the array of preclinical cognitive and dementia-relevant outcome measures on which our preventive interventions would be expected to have an impact. The focus is on measures that can pick up subtle change in cognition and behavior in early to midlife and that are known to be linked to the distal outcome of AD/ADRD. Measures would need to capture stability, improvement, or slowed decline in cognitive function, to indicate that our intervention is working as hypothesized. Evidence for the links between these outcome measures and the target processes (in Panel 1), and suitability for their inclusion as outcome measures in prevention trials will be discussed. What do we need to know to be confident that these measures serve as appropriate indicators of progression or non-progression to AD/ADRD?

3:00 pm Introduction – Richard Gershon

3:05 pm Standard Neuropsychology as an Outcome: Challenges and Opportunities – Yakeel Quiroz

3:25 pm Break

3:30 pm Remote Cognitive Measurement as an Outcome – Richard Gershon

3:50 pm Found Measures of Cognition – Lauren Nicholas

4:10 pm Discussion: What more do we need to know about measuring dementia-relevant cognitive outcomes in midlife?

4:30 pm Wrap up, Day 1

Day 2: May 25, 2021

1:00 pm Welcome to Day 2 – Jonathan W. King

Panel 3: Issues in Intervention Development: How to Design the Most Informative Trials

Chair: Linda M. Collins

This Panel will explore how we can generate causal evidence through experimentation, before launching large-scale prevention trials, thereby creating greater likelihood of success and improved understanding of tailoring needs. The use of experimental strategies, MOST designs, and identifying appropriate populations based on risk factors is emphasized.

1:05 pm Introduction – Linda M. Collins

1:10 pm	Developing Effective, Affordable, Scalable, and Efficient Interventions – Linda M. Collins
1:30 pm	Who Do We Really End Up Recruiting into Our Trials and Research? – Melinda C. Power
1:50 pm	Large Trials in the Health Care System to Detect and Prevent Dementia – Jason Doctor
2:10 pm	Break
2:15 pm	Discussion

Panel 4: Non-intervention Development: Generating Causal Evidence from Observational and Non-clinical Studies to Inform the Prevention Research Agenda

Chair: Maria Glymour

Given the logistical difficulties of running long-term prevention trials beginning in midlife but simultaneously recognizing advances in causal modeling from observational data, we might also reconsider how best to generate evidence on prevention from these studies. This panel will review some of the overall approaches and particular sources of data we might use to strengthen the evidence base for treatment effectiveness or efficacy.

2:35 pm	Introduction – Maria Glymour
2:40 pm	Causal Inference from Observational Data – Sonja Swanson
3:00 pm	Quasi-Experimental Methods – Emilie Courtin
3:20 pm	Break
3:25 pm	Measuring Effects of Treatment Variation in Healthcare Systems and Their Effects on Cognition – Rachel Whitmer
3:45 pm	Chronic Conditions and Risk of AD/ADRD: Advancing Prevention Analyses to Inform Population Health – Julie Zissimopoulos
4:05 pm	Discussion: What can we do to increase confidence in observational results as evidence of treatment efficacy or effectiveness?
4:25 pm	Break
4:30 pm	<i>General Discussion</i>
5:00 pm	End of Meeting