Table of Contents

Executive Summary ................................................................................................................. 1
Improving and Expanding RWD Infrastructure ...................................................................... 1
Improving RCTs While Supporting RCT Alternatives .............................................................. 2
Improving Demographic, SDOH, and Health Disparity Analyses ........................................... 2
Ethical, Privacy, and Regulatory Considerations .......................................................................... 3

Meeting Summary .................................................................................................................... 4
Welcome Remarks and Charge—Richard Hodes .................................................................. 4

Session 1: Overview of Current Infrastructure ......................................................................... 4
Gaps and Opportunities in Current Real-World Data Infrastructure – Atul Butte .................... 4
RWD Infrastructure, the Alzheimer’s Association Network Registry, and Opportunities for Improved Patient Outcomes – Maria Carrillo ................................................................. 5
Real-World Evidence in Alzheimer’s Disease: The ROADMAP Data Cube – Olin Janssen .... 7
Key Barriers to Clinical Trials for Alzheimer’s Disease – Niranjan Bose ......................... 8
Panel Discussion ....................................................................................................................... 10

Session 2: Recruitment and Clinical Trial Uses of Real-World Data ....................................... 11
Overview of Possibilities of Real-World Data for Clinical Trials and Patient Experience – Adrian F. Hernandez ..................................................................................................................................... 11
Race and Ethnicity in Real-World Data Sources – Michelle E. Tarver .................................... 12
Recruitment for Alzheimer’s Disease Prevention Trials – Rema Raman ................................ 13
Recruitment for Dementia Care Trials – Vincent Mor .............................................................. 14
Panel Discussion ....................................................................................................................... 14

Session 3: Secondary Data Uses of Real-World Data ................................................................ 16
Which Scientific Effectiveness Questions can be Answered with Data from Clinical Practice? – Sebastian Schneeweiss ................................................................................................. 16
RWD Linkages: Possibilities for NIA Longitudinal Studies – Maria Glymour ....................... 17
Using RWD for Identifying and Validating Drug Repurposing Candidates for AD/ADRD – Mark W. Albers ............................................................................................................................. 18
Unification of Clinical Trial Data – Gabriel S. Eichler .............................................................. 19
Panel Discussion ....................................................................................................................... 19

Session 4: Discussion on Infrastructure Gaps and Opportunities ............................................. 22
Invited Comments ...................................................................................................................... 22
Invited Commenter Discussion: Building Equitable Real-World Data Infrastructure that Addresses Wide-Ranging Stakeholder Needs .................................................................................. 25

Closing Panel: Prioritizing Gaps and Opportunities for Using RWD in AD/ADRD Research .......... 26
Marina Sirota .......................................................................................................................... 26
Malaz Boustani ...................................................................................................................... 27
David Dore ............................................................................................................................. 27
Kenneth Langa ....................................................................................................................... 28
Julie Zissimopoulos .................................................................................................................. 28

Appendix 1. Workshop Participants ......................................................................................... 29
Presenters ................................................................................................................................. 29
NIA RWD Working Group (*Planning Committee Members) ................................................... 30

Appendix 2. Agenda .................................................................................................................... 31
Executive Summary

On May 4, 2022, the National Institute on Aging (NIA) convened a panel of subject matter experts to explore gaps and opportunities for real-world data (RWD) infrastructure that supports research and clinical trials for Alzheimer’s Disease and Alzheimer’s Disease and Related Dementias (AD/ADRD).

Panelists highlighted limitations in the current infrastructures, including challenges in integrating disparate data sources, ensuring data validity and generalizability, and facilitating data deidentification without losing valuable information. Panelists also discussed opportunities for RWD to improve recruitment and retention in randomized control trials (RCTs) and pragmatic trials, assess the efficacy/effectiveness of drug repurposing, identify new therapeutics, link data to NIA-funded longitudinal panel studies, and improve regulatory decision-making for novel AD/ADRD therapeutics. Panelists noted how RWD can enhance research on the social determinants of health (SDOH) and health disparities by providing researchers a larger sampling frame from which to enroll diverse study or trial participants.

Presenters and panelists emphasized the need for a patient-centered approach when establishing or enhancing RWD infrastructures, which must integrate ethical and privacy concerns (including appropriately consenting participants), at both the data collection and data analysis stages of RWD-based research. Of particular importance the potential for misuse of RWD and real-world evidence (RWE), the potential for investigators to inadvertently perpetuate biases that often exist in analytical algorithms or in the data themselves when they are transformed, and ethical concerns related to patients with AD/ADRD. RWD infrastructures for AD/ADRD research must serve and include the communities who would most benefit from the treatments, innovations, and ideas that emerge from that infrastructure.

Key themes from the workshop included:

**Improving and Expanding RWD Infrastructure**

RWD currently exist in many different siloes, including electronic health records (EHRs), patient and physician registries, claims databases and many more. Integrating and harmonizing these disparate data sources offers opportunities to expand RWD analyses. Incorporating data sources such as data from the U.S. Census Bureau and Social Security Administration (SSA) in a RWD infrastructure could provide additional information on social determinants of health (SDOH), financial impacts of AD/ADRD on patients and caregivers, and health disparities. One possible approach is establishing a modernized, next-generation health-information exchange that is optimized for both improving health care and RWD research. However, integrating these data sources requires overcoming multiple hurdles, including the need to handle disparate data elements, the difficulties of extracting relevant data from unstructured data sources (e.g., clinician notes), and deidentifying data without losing necessary information.
Research Gaps and Opportunities

1. Develop common data elements – or otherwise harmonized data – to ensure greater interoperability between RWD sources
2. Improve access to RWD
3. Develop new deidentification models that preserve valuable research information while complying with The Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirements, specifically the HIPPA Privacy Rule

Improving RCTs While Supporting RCT Alternatives

RWD provide agility to clinical trials (e.g., via passive collection of study data) and support an array of secondary data analyses. For clinical trials, RWD can improve recruitment and representativeness through precision recruitment, streamlining screening criteria, and identifying barriers to trial participation. Pragmatic or decentralized trials can also leverage RWD to improve sampling frame size and diversity. Furthermore, not all scientific questions can be answered by RCTs given cost limitations, ethical concerns, or pragmatic limitations. In these instances, RWD can be leveraged to conduct trial emulations, or other forms of secondary data analyses of drug effectiveness, drug repurposing, and long-term analyses of AD/ADRD risk factors, resilience strategies, and changes in symptom presentation throughout the AD spectrum.

Research Gaps and Opportunities

1. Promote the use of RWD to increase clinical trial recruitment speed and participant diversity
2. Employ flexible and decentralized clinical trial designs that incorporate virtual, at-home, and point-of-care approaches to reach more participants, particularly in areas far from clinical trial sites
3. Support secondary analyses of RWD that can answer scientific questions about drug repurposing, effectiveness, and long-term outcomes.
4. Incorporate Agile frameworks, network science, and behavioral economics to improve clinical trial designs and recruitment

Improving Demographic, SDOH, and Health Disparity Analyses

RWD provides opportunities to link information on patient heterogeneity to research on demographic or within-group differences in health outcomes. Linking health data (e.g., EHRs, claims data, death indices, patient registries) with geographic information (e.g., ZIP code, Area Deprivation Index [ADI] scores) can highlight geographic and socioeconomic disparities, including in areas where community-based approaches can significantly improve access to care. Improving collection and integration of demographic data can also support the generation of valid RWE that translates to diverse populations. Lastly, RWE studies should use a strengths-based approach to addressing health equity, and involve the communities being studied in the research process from start to finish.
**Research Gaps and Opportunities**

1. Integrate many different data sources (e.g., U.S. Census Bureau data) into RWD infrastructures
2. Standardize processes for collecting racial and demographic information, including the ability for patients to select multiple categories
3. Account for intersecting identities (e.g., race, gender, sexual orientation) when analyzing demographic differences in health outcomes and access to care
4. Incorporate cultural and community perspectives into research using RWD

**Ethical, Privacy, and Regulatory Considerations**

While expanded RWD infrastructure and analyses offer unprecedented opportunities for advancing AD/ADRD research and treatments, these analyses and infrastructures must consider potential ethical, privacy, and consent risks. Many minoritized groups are underrepresented or misrepresented in EHRs, registries, and other RWD sources, and biases in underlying data can cause RWD analyses to inadvertently perpetuate biases and stereotypes regarding marginalized groups—limiting the generalizability of results. As data sources are increasingly integrated, researchers and data scientists must consider the potential unintended consequences of using misguided methods when accessing certain RWD sources. Continually adding more deidentified data tied to the same individual may also increase the risk of reidentification, requiring more rigorous data security practices. Furthermore, while RWD offers multiple opportunities for large decentralized pragmatic trials, data safety monitoring boards (DSMBs) and institutional review boards (IRBs) are still adapting procedures for managing and monitoring decentralized clinical trials and pragmatic trials driven by RWD.

**Research Gaps and Opportunities**

1. Engage multiple stakeholder perspectives throughout the research process and share results and conclusions generated from RWD with the individuals and community who provided the data
2. Develop updated consent and ethics frameworks that apply to decentralized RCTs and pragmatic trials
3. Address reluctance among minorities with historic negative experiences with medical research through greater transparency and community engagement
Meeting Summary

Welcome Remarks and Charge—Richard Hodes
Dr. Hodes introduced the workshop, which aimed to identify gaps and opportunities in current real-world data (RWD) infrastructures, and discuss current or potential capabilities for using RWD to support research on aging, Alzheimer’s disease (AD), and Alzheimer’s disease and related dementias (ADRD).

Session 1: Overview of Current Infrastructure

Gaps and Opportunities in Current Real-World Data Infrastructure – Atul Butte
Given the billions of dollars being spent on implementing, maintaining, and improving electronic health record (EHR) systems in the United States, a failure to use these data to improve medical practice would be a national tragedy. The University of California (UC) Health System (UC Health) has developed multiple techniques to enable analyses of pooled real-world data (RWD) to improve medical care.

UC Health includes 10 campuses, three national laboratories, 20 health professional schools (including six medical schools), and five National Cancer Institute (NCI)-designated Comprehensive Cancer Centers. UC Health was established as a way for the six UC academic health systems to leverage their scale in health system operations and research to create an accountable care organization (ACO). To reach this goal, data from the EHR systems used in the main hospitals and clinics across UC Health are pooled into a central data warehouse and data analytics platform. The UC Health central data warehouse uses the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) to harmonize disparate data types (e.g., diagnosis codes, laboratory test results, medication orders) to support longitudinal analyses of health outcomes, and UC Health leverages these analyses to examine and improve quality of care, treatment outcomes, and cost savings. UC Health has integrated California regulatory data, including mortality, pathology, and radiology text data, into this centralized system, and now integrates the California state death index.

UC Health also collaborates with Observational Health Data Sciences and Informatics (OHDSI), an international network of health researchers and federated databases that pool data and share analytical techniques to improve sample sizes and data quality for clinical research. OHDSI also provides scorecards that measure contributor data quality, on which UC Health scores highly.

As an example of the quantity of data within the UC Health Data Warehouse, Dr. Butte showed a distribution of the 24 million serum potassium levels measured in UC Health patients over the past 10 years. He also showed how RWD collected on drugs post-approval can provide more datapoints than the original pivotal studies designed to test drug effectiveness and other outcomes. For example, the original pivotal clinical trials of Humira had less than 900 total participants per study, whereas UC Health now has data on more than 11,000 patients using
Humira. This growing dataset can reveal many insights, including novel side effects and real-world effectiveness, on drugs, biologics, medical devices, and other therapeutics.

UC Health has identified 21 uses for RWD, broadly including (1) supporting regulatory approvals; (2) examining post-approval safety of therapeutics, including side effect rates; (3) informing clinical trial design; (4) establishing efficacy of therapeutics and treatments; (5) assessing comparative effectiveness; (6) studying medical practices, including medical error rates, effects of payors on care decisions, and diagnostic effectiveness; and (7) supporting data-driven health care decisions.¹

Pooled UC Health patient data are initially fully identified, and include ZIP codes, race, gender, and address-based calculated indices for social determinants of health (SDOH). Identifiable data is used for operational improvements, and documenting and enhancing the quality of delivered patient care. De-identified data are available for research to UC staff, faculty, students, or post-docs via single sign-on through a secure cloud-based system, and a virtual machine enables researchers to analyze relevant data from the data warehouse in Azure using the Databricks platform. Extracted data cannot be downloaded. Before access, researchers have to sign an attestation demonstrating understanding of the terms and rules of access. This system provides safeguards to protect privacy while also increasing data accessibility.

To further increase the amount and types of RWD available, UC San Francisco (UCSF) developed an open-source algorithm that removes identifiable information from the unstructured data in clinical notes. This targeted removal prevents the unintended loss of valuable unidentifiable details. Using this algorithm, UCSF de-identified health data from over 110 million EHR notes, increasing availability of these data for research. UC Health collaborates with institutional review boards (IRB) and the UC Information Security and Privacy departments to ensure data are fully de-identified. This methodology is now being considered for scaling across the rest of the University of California Health System.

The Food and Drug Administration (FDA) provides extensive resources and frameworks to guide researchers on using RWD for a variety of use cases, including drug approvals without clinical trials. Partnerships between FDA and NIA may be important moving forward, given the wealth of knowledge FDA has already produced on RWD.

RWD Infrastructure, the Alzheimer’s Association Network Registry, and Opportunities for Improved Patient Outcomes – Maria Carrillo

Advances in AD biology and genetics research, along with the development and approval of monoclonal antibodies (mAbs) that target amyloid and tau protein accumulation, offer the potential for a new era of AD diagnosis and treatment. However, recent safety concerns over aducanumab and other amyloid-targeting mAbs highlight the need for additional safety and effectiveness data for AD therapeutics, particularly in real-world environments.

In November 2021, the Alzheimer’s Association announced the development of the Alzheimer’s Network for Treatment and Diagnostics (ALZ-NET), a national physician-based registry that collects RWD from clinical practice through the Network on the usage, safety, and effectiveness of novel FDA-approved AD therapeutics. ALZ-NET aims to build upon successes from other disease areas that utilize RWE collection to understand and address the evolving standards of care and novel treatment approaches.

ALZ-NET will leverage the demographically diverse infrastructure developed through the Imaging Dementia – Evidence for Amyloid Scanning (iDEAS) study, which includes more than 1,000 dementia experts, 500 dementia clinics, and 300 positron emission tomography (PET) facilities. iDEAS was a single-arm, multi-site, longitudinal study that demonstrated the effectiveness of using amyloid PET imaging for early AD diagnosis.

ALZ-NET has multiple objectives:

- Develop a multi-site network for enrollment and data collection.
- Build and implement resources for the clinical readiness in a new era of treatment (i.e. trainings).
- Collect longitudinal patient data from the start of treatment that includes measures of cognition, function, and patient safety.
- Track participant health outcomes and resource utilization.
- Collect and archive brain scans, blood/biofluid samples, and genetic and biomarker data.
- Share de-identified data, brain scans, and blood/biofluid samples with the research community to support continuing AD research.

The ALZ-NET team is currently establishing a network infrastructure to gather regulatory grade, longitudinal data from patients being treated with novel FDA-approved AD therapeutics. ALZ-NET is also establishing an image repository for diagnostic and neuroimaging studies (including magnetic resonance imaging and PET) that incorporates cognition and function assessments, adverse event (AE) and serious adverse event (SAE) reporting, patient demographics, and biomarker results. ALZ-NET will also establish an infrastructure for sharing de-identified data and images, as well as a biorepository for collecting and archiving biosamples from patients consented in ALZ-NET. Once ALZ-NET is established, researchers will be able to use the network to evaluate longitudinal safety and health outcomes, including cognitive changes, safety measures, and physician prescribing patterns.

ALZ-NET will begin tracking patients when they initiate treatment with novel AD therapeutics and will incorporate future therapeutics as they are approved by FDA. ALZ-NET aims to evolve with changes in AD diagnostics and treatments and to use RWD to continuously improve safety monitoring of novel AD therapeutics.
Real-World Evidence in Alzheimer’s Disease: The ROADMAP Data Cube – Olin Janssen

The Real-world Outcomes across the AD spectrum for better care: Multimodal data Access Platform (ROADMAP) project was a two-year public-private partnership (PPP) managed by the European Federation of Pharmaceutical Industries and Associations (EFPIA) Innovative Medicines Initiative (IMI) to assess the current state of different AD RWD sources and their usability for developing disease models, treatments, and health policy.

RWD for AD span a wide range of sources that are applicable to different stages of the disease. For example, cohort studies of amyloid biomarkers and memory clinical studies can provide insights into preclinical AD and mild dementia, whereas EHR data are often limited to more severe disease stages. Relevant health and quality of life measures also differ as the disease progresses, and the relative importance of measures may differ across stakeholders, such as AD patients, caregivers, and health care providers (HCPs). Thus, examining the interactions of data sources, disease stage, and stakeholder type can help researchers identify the appropriate data source to examine for any given use case, depending on types of information sought (e.g., changes in cognitive ability across disease stages) as well as information gaps in that data source (e.g., limited data on caregiver quality of life). No single data source includes outcomes that span all stages of the AD spectrum.

The ROADMAP project built upon previous AD RWD projects (e.g., Dementias Platform United Kingdom [DPUK]) that provided insights into characteristics of different data sources in relation to various disease outcomes across the AD spectrum. To identify AD outcomes of importance to various stakeholders, ROADMAP conducted a review of relevant AD literature, held a series of patient and public consultations, and reviewed results from stakeholder surveys that were informed by key outcomes from the literature review and consultations.

Key survey findings included:

- All stakeholder groups considered patient quality of life and significant disease-related life events as the most important measures across all disease stages.
- Most respondents rated cognitive ability, functional ability, and independence as important during the Mild Cognitive Impairment (MCI) stage, but the importance of these constructs decreased at later disease stages.
- At later disease stages, most respondents rated behavioral and neuropsychiatric symptom management and quality of caregivers’ lives as most important.

ROADMAP researchers integrated these data into the ROADMAP data cube, which aims to provide high-level information about data availability through a three-dimensional heat map based on three dimensions: data source, disease stage, and AD outcome measurements. This data cube can be filtered by stakeholder type, disease stage, and data type; darker colors indicate greater data availability.

ROADMAP researchers also conducted a gap analysis of data availability based on a subset of 66 data sources from the data cube, including data from four clinical trial placebo arms, four EHR databases, and 58 cohort studies. While more than 80 percent of these data sources
contained data on cognitive ability, comorbidities, and therapeutic treatment types, less than 30 percent contained data on neuroimaging biomarkers, caregiver use of health and social services, and caregiver quality of life. Clinical trial placebo arms contained all measures except caregiver quality of life, while EHRs did not include cognitive ability, neuroimaging biomarkers, APOE-e4 scores, or caregiver use of health and social services. Less than 25 percent of cohort studies included data on caregiver quality of life, use of health and social services, or neuroimaging biomarkers.

Multiple data sources are needed to capture all relevant outcomes across the AD spectrum. Therefore, researchers should consider the suitability and limitations of different data sources when using RWD. ROADMAP’s limitations include a lack of patient-specific data and limited EHR and clinical trial data on outcomes in severe AD. Improving the quality and usability of RWD for AD requires worldwide collaboration, a standard data platform with appropriate data governance and harmonization processes, the ability to share de-identified datasets with researchers, and engagement with AD patients.

**Key Barriers to Clinical Trials for Alzheimer’s Disease – Niranjan Bose**

AD clinical trials continue to face logistical challenges, such as long trial duration, slow enrollment, and high costs. Compared to other therapeutic areas, disease modifying therapy (DMT) trials for AD have slower enrollment rates, higher screening failure rates, and longer trial durations. Furthermore, AD DMT trials cost more per patient than most trials in other therapeutic areas due to higher costs for patient screening and randomization, which can account for 50 to 70 percent of total per-patient trial costs. This higher cost is driven by high screening failure rates associated with early AD, and a significant loss to follow-up early in patients’ journey through the clinical trial ecosystem.

The patient journey into clinical trials begins in the public ecosystem, in which most patients are relatively disengaged from the healthcare system. When AD symptoms develop, some patients self-refer to clinical trials, while others are referred through their general practitioner, an AD or brain health specialist, or an AD registry. Challenges in this recruitment process were examined by IQVIA researchers, who identified eight key barriers to recruitment:

- Limited patient awareness of AD, including misconceptions about AD symptoms being “normal aging.”
- Fear and stigma around AD, causing many patients to deny symptoms or diagnosis.
- Overstretched health care systems, including long wait times for specialists.
- Poor physician awareness of early AD symptoms, particularly preclinical and MCI symptoms.
- Lack of fast and inexpensive diagnostics.
- Poor perception of AD treatment options.
- Awareness of and referral to AD clinical trials.
- Limited diversity of clinical trial subjects.
Researchers conducted multiple workshops with key stakeholders to identify potential solutions to clinical trial recruitment, and these solutions were graphed by potential impact and level of effort/investment. Key high-impact solutions included establishing blood-based biomarker tests for AD, centralized data-sharing platforms (which can integrate EHRs and registries), and Centers for Medicare and Medicaid Services (CMS) reimbursement for diagnostics.

USC Schaeffer, in collaboration with the Alzheimer's Therapeutic Research Institute and Howard University, are launching multiple pilot innovations to address many AD clinical trial challenges. Pilot projects include (1) a feasibility assessment of remote collection of plasma biomarkers, (2) new recruitment models using community-based partnerships, and (3) a mobile self-administered cognitive assessment tool. These pilot projects are guided by robust measurement plans, evidence-based approaches, and feedback and guidance from an expert advisory panel.
Panel Discussion

**Moderator** – Jennie Larkin

**David Dore**
RWD infrastructures still face challenges in incorporating EHR data into a standard architecture such as OMOP, and in extracting the most complete data possible from unstructured data sources while remaining compliant with ethical standards and privacy law. Data collection efforts for registries and clinical trials must consider the burden associated with primary data collection; collecting secondary data, such as EHRs, can reduce this burden. When developing or using a registry, researchers must also consider the impact of enrollment requirements on the representativeness of the study population. Currently, there is increasing emphasis placed on moving registries and clinical trials closer to the site of care to streamline clinical trial participation, improve recruitment, and increase representativeness—as in the NIA-funded IMPACT Collaboratory.

**Barbara Entwisle**
Although researchers face several barriers to using U.S. Census Bureau data, these data can enhance and inform the use of other RWD sources. For example, Census data can add value to EHRs, which often lack information on individual and household characteristics. Census data can be used to determine the representativeness of other RWD sources and to provide greater context for the consequences of various medical outcomes on individuals and households—for example, how a dementia diagnosis impacts each member of a household.

**Stephanie Monroe**
Including underserved communities in RWD sources is valuable for establishing RWD quality and accuracy that can facilitate the development of meaningful RWE. If unaddressed, the use of biased RWD sources can lead to dangerous misapplications of data and incorrect conclusions that can exacerbate the very problems RWD infrastructures intend to solve. This risk is particularly problematic given that underrepresented populations are likely to be the primary consumers of AD/ADRD treatments developed through RWE. Training can help improve accurate and objective data collection by healthcare staff. For example, African American patients are twice as likely to be described negatively (e.g., non-compliant, aggressive, etc.) in clinician notes compared to non-Hispanic white individuals. Lastly, infrastructures must be built in collaboration with the communities they intend to serve.

**General Discussion**
RWD infrastructures offer great room for innovation. Dr. Dore noted that biases in RWD and RWE are propagated when researchers do not select appropriate design and analytical methods; collaborations across multidisciplinary groups that include expertise in correcting for underrepresentation or misrepresentation in datasets is crucial for addressing this problem. Dr. Entwisle emphasized that representativeness and bias in the data themselves must also be considered. She further noted that datasets often do not account for within-group diversity.

Dr. Monroe added that zip code data, which are often removed during deidentification, are useful to collect for examining SDOH and associations between various health outcomes and
the lived environments people inhabit. Dr. Dore noted that researchers need not be anchored to currently existing deidentified data models and suggested that variations in deidentified data models can allow for inclusion of additional information (e.g., geographic data) if other types of potentially identifying data are removed. Lastly, communities that consent to provide their data should benefit from the results produced by the data analysis.

**Session 2: Recruitment and Clinical Trial Uses of Real-World Data**

**Overview of Possibilities of Real-World Data for Clinical Trials and Patient Experience – Adrian F. Hernandez**

Only approximately two percent of the U.S. population participates in clinical trials, while survey data show that at least 85 percent of Americans are interested in participation. Clinical trials that enroll larger sample sizes and more representative patient populations would provide more nuanced and generalizable data, better support pragmatic and efficacy trials, and provide additional RWD for assessing real-world efficacy of different treatments. Key barriers to clinical trial participation include the widespread existence of clinical trial deserts (i.e., areas of the country far from clinical trial sites) and limited broadband internet access in many rural areas, which hinders participation in virtual trials and follow-ups.

RWD can be leveraged to improve clinical trials by (1) streamlining clinical trial operations, (2) reducing participant burden, (3) improving data quality by integrating RWD data sources with clinical trials, and (4) validating clinical trial design, data quality, and generalizability of results. Clinical trials can also be improved by incorporating pragmatic approaches, such as embedding research into clinical care, using patient-reported data, and capitalizing on advances in remote clinical trials and follow-ups. Furthermore, implementing a common data model (CDM) such as PCORnet can enable integration of multiple RWD sources across different platforms and data siloes.

The potential of RWD and pragmatic clinical trial designs to improve clinical trial participation was demonstrated in the PRagmatic Evaluation of evENTS And Benefits of Lipid-Lowering in oldEr adults (PREVENTABLE) trial. PREVENTABLE examined the potential of statin medications to prevent dementia, disabilities, and heart disease in adults over 75. PREVENTABLE eased participant burden by shipping study medication directly to participants’ homes, collecting data online and through EHRs, and using virtual and telephone-based consent and recruitment processes. Similarly, the Accelerating COVID-19 Therapeutic Interventions and Vaccines (ACTIV)-6 trial employed a hybrid approach that combined virtual and site-based clinical trial approaches to study repurposed drugs (e.g., ivermectin, fluticasone) for treating COVID-19 symptoms in non-hospitalized adults. ACTIV-6, which had more than 15,000 participants across all 50 states, was managed using a central online participant portal, used a central call center for services, and directed drug shipments to participants’ homes. Both PREVENTABLE and ACTIV-6 demonstrate the ability of pragmatic trial designs to overcome many challenges associated with traditional site-based clinical trials.
Race and Ethnicity in Real-World Data Sources — Michelle E. Tarver

Capturing and analyzing demographic differences in RWD is crucial given long-standing health disparities between demographic groups, particularly racial and ethnic groups. The COVID-19 pandemic revealed stark health disparities, with older populations and racial and ethnic minorities more likely to experience worse health outcomes from SARS-CoV-2 than their younger and Caucasian counterparts. Furthermore, differences in biology and environment can impact responses to disease and medical products. For example, women are more likely than men to have a stronger inflammatory response to certain medical devices.

Capturing racial and ethnic demographic data requires understanding the differences between race, ethnicity, and ancestry:

- Race is a sociocultural concept recognized in the United States, and may not reflect biological, anthropological, or genetic differences.
- Ethnicity refers to the shared social, cultural, and historical experiences stemming from common heritage, nationality, lineage, and country or region of birth.
- Ancestry refers to the ethnic origin, descent, heritage, or place of birth of a person or their ancestors.

Racial and ethnic data should be collected and considered throughout the medical product lifecycle, including during development, clinical trial design, and RWD collection. Both FDA and the Office of Management and Budget (OMB) provide minimum recommendations for race and ethnicity demographic data collection to facilitate and standardize studies on medical products.

Although RWD sources can provide large amounts of information on health outcomes across racial and ethnic groups, many data sources have missing or inaccurate racial and ethnic data. 11 to 22 percent of hospital systems are estimated to not collect any racial or ethnic data. 17 percent of patients report discomfort disclosing their race or ethnicity, and 43 percent of these patients express concern about their racial or ethnic data being used to discriminate against them.² Some racial and ethnic groups—particularly Hispanic, Middle Eastern, and Native American patients—are misclassified as white, particularly when race is recorded by someone other than the patient.³ These data gaps and misclassifications hinder efforts to integrate racial and ethnic data across different RWD sources and to accurately analyze differences in health outcomes between racial and ethnic groups.


Meeting Summary
Addressing these challenges requires the use of consistent and appropriate racial and ethnic terminology across health care systems and RWD sources, with greater involvement of patients and underrepresented groups in the development of systems for capturing such data. Furthermore, a standard approach is needed to mitigate differences that arise from variation in local regulations about the collection and analysis of racial and ethnic data. Finally, greater transparency and community involvement can help mitigate potential unanticipated or discriminatory consequences of collecting racial and ethnic data.

**Recruitment for Alzheimer’s Disease Prevention Trials – Rema Raman**

External validity is crucial for AD/ADRD clinical trials. However, many trials fail to capture participant populations that match the target population, limiting the generalizability and applicability of the results. Many demographic groups are underrepresented in AD/ADRD trials, including racial and ethnic minorities, LGBTQ+ people, individuals with a lower socioeconomic status (SES), and individuals living in geographic regions far from clinical trial sites. Increasing clinical trial diversity and external validity requires improved strategies for recruiting and retaining a diverse and representative pool of participants throughout the clinical trial process, from initial recruitment and prescreening through trial completion.

The Alzheimer’s Clinical Trials Consortium (ACTC) provides infrastructure to accelerate the development of effective interventions for AD/ADRD. The ACTC Recruitment Unit is a site- and participant-focused, culturally sensitive infrastructure that promotes data-driven and evidence-based recruitment and retention. This infrastructure includes:

- **Outreach data** from website pre-screening forms, site-level outreach, and central recruitment activities.
- **Pre-screening data** on race/ethnicity and reasons for not enrolling.
- The **recruitment minimal data set (MDS)**, which encompasses much more detailed information, including demographics, reasons for participants being screened out, Area Deprivation Index (ADI), and reasons for enrolling. These data enable detailed analyses of participants, including factors that may lead to certain populations’ underrepresentation.

ACTC, in collaboration with NIA and the Alzheimer’s Association, is also training the next generation of AD/ADRD clinical trial researchers as part of the Institute on Methods and Protocols for Advancement of Clinical Trials in ADRD (IMPACT-AD) program.

RWD’s limitations must also be recognized when applying RWD to clinical trial recruitment and analysis. However, RWD, particularly after effective treatments become available, can be used to potentially increase representativeness of clinical trial participants in AD/ADRD clinical trials; RWD enable researchers to use direct campaigns to underrepresented groups, design eligibility criteria based on the target population, and identify solutions for populations with inadequate access to clinical trial sites. These approaches will require greater collaboration among cross-functional teams, including clinicians, biostatisticians, data scientists, and bioinformatics experts.
**Recruitment for Dementia Care Trials – Vincent Mor**

The NIA IMPACT Collaboratory aims to build the nation’s capacity to conduct pragmatic clinical trials of interventions embedded within health care systems for individuals living with AD/ADRD. IMPACT leverages RWD through two research networks. The first is CVS Clinical Trials Services, which IMPACT uses to match pharmacy customer data to Medicare claims data to conduct observational analyses. IMPACT has also leveraged CVS’s large customer base to increase trial participation through precision patient recruitment. For example, the IMPACT-CVS Protocol Development Project focused recruitment on store customers in high minority areas to increase trial diversity. IMPACT’s use of CVS Clinical Trials Services has increased opportunities for decentralized trials and for RWD generation through the assessment of interventions in real-world settings, such as “Minute-Clinics.”

The second network is the Long Term Care (LTC) Data Cooperative, which has assembled electronic medical record (EMR) data from thousands of nursing homes; these data are linked to Medicare claims, which facilitates the identification of eligible study subjects for Phase III and IV trials. The LTC Cooperative includes a subset of nursing home companies willing to participate in select treatment trials. IMPACT used the LTC Data Cooperative to conduct a study examining the effects of an intervention on encouraging COVID-19 vaccination among skilled nursing facility (SNF) residents and staff at 133 randomized SNFs across 16 states. This study included more than 7,000 SNF residents and 15,000 SNF staff equally split between intervention and control arms. Although the study did not show an effect of the intervention on vaccination status, it serves as a proof of concept for large-scale embedded trials using RWD to identify subjects and monitor outcomes.

Using RWD for pragmatic trial recruitment offers numerous benefits, including ease of participant identification and the ability to identify caregivers and increase representativeness in trials through precision recruitment in high minority areas. However, Dr. Mor cautioned that inadequate diagnostic accuracy in RWD must be considered.

RWD can benefit clinical trial recruitment and analysis, although limitations must always be recognized. Primary data collection during Phase III trials can bolster the use of RWD in conducting large-scale pragmatic interventions in Phase IV. Random assignment can aid internal validity of results, and the limitations of using RWD are offset by the potential for recruiting more representative populations than most efficacy or effectiveness studies.

**Panel Discussion**

*Moderator – Marcel Salive*

**Malaz Boustani**

Dr. Boustani emphasized that speed and agility are the paramount advantages of RWD, and these unique qualities must be leveraged in clinical trial design. Advances in Agile science, behavioral economics, and network science (i.e., using a “building block” approach) can improve clinical trial recruitment and retention with the goal of transitioning “strangers” of research recruitment into “fans.” These advances can also be used to determine how individuals can be clustered to enhance the speed of discovery and can offer a streamlined
approach by which researchers can personalize trials to specific communities. Lessons learned from COVID-19, such as the value of home-based recruitment, can be leveraged moving forward.

In response to the COVID-19 pandemic, Indiana University created a research recruitment and retention nudge unit, which mapped the journey of individuals from “strangers” to “fans” of trial recruitment. This process identified three key stages of patient contact. During the first stage, digital contact incites interest among potential participants. The messenger selected to deliver these initial contacts is critical. During the second stage, relationship-building replaces digital outreach; trust, time, and space are integral for fostering these relationships. The third stage, termed the “contract,” requires an efficient consent and follow-up process to streamline recruitment. Transitioning new recruits into “fans” who may even recruit other participants occurs through managing a well-run trial or study and sharing study results with participants after the study’s conclusion.

Emily Largent
Dr. Largent noted the tension between equity and pragmatism, but emphasized that trials not designed to address health disparities will perpetuate health disparities. Disparities exist across many dimensions, including geography (e.g., urban/rural), race/ethnicity, socio-economic status, and sex/gender.

The collection and use of RWD for pragmatic and decentralized clinical trials poses unique ethical challenges, particularly for consent. While waivers are often used in decentralized trials, participants must still be notified when research is occurring and be educated on the importance of large real-world trials. Certain groups—particularly those who have experienced stigma and discrimination within the health care system—have (when interviewed by researchers) expressed hesitation to participate in research conducted with a waiver of informed consent. Providing education and establishing trust are key ways by which researchers can address this hesitation to participate. Another challenge in recruiting individuals, especially for dementia research, is ensuring that researchers do not inadvertently disclose a diagnosis to an unknowing participant during a “cold call” or other procedure when recruiting participants to trials.

Lastly, ethical and regulatory protections for participants must evolve to emerging challenges for trials using RWD. IRBs and trial data safety monitoring boards (DSMBs) are still defining roles and procedures for managing and monitoring decentralized clinical trials. Similarly, guidance about investigator and institutional obligations to pragmatic trial participants is still lacking. Traditional models cannot be simply transferred from traditional trials to pragmatic trials.

Corrie Painter
Many themes from this workshop are prevalent across other healthcare fields, and lessons can be gleaned from other fields’ experiences in these areas. In Dr. Painter’s work with cancer patients, data collection vehicles combine patient-reported data, survey data, tissue samples, and germ-line samples to provide open-source aggregate data for researchers to study. This
effort demonstrates that a patient-centered approach can be conducted at-scale. Additionally, data in RWD infrastructures must also be as representative as possible, and Dr. Painter’s data collection vehicles aim to improve representation by implementing a Community Engagement Team. Lessons learned from these data collection efforts likely translate to other disease types, including AD/ADRD.

Lastly, many RWD sources were not originally designed for RWD research purposes, which presents ethical concerns about using such data without involving patients and their caregivers. Integrating patients and caregivers into RWD collection and research can improve AD/ADRD trials and studies.

**Session 3: Secondary Data Uses of Real-World Data**

**Which Scientific Effectiveness Questions can be Answered with Data from Clinical Practice? – Sebastian Schneeweiss**

Complementing randomized clinical trials (RCTs) with RWE analyses can improve the breadth (e.g., different patient subgroups), depth (e.g., different effectiveness and safety endpoints), and clinical context (e.g., different comparators and treatment combinations) of available medical information. Although RCTs are immensely useful, they cannot answer many scientific questions due to ethical concerns and cost. Many of these unanswered questions can be examined through database studies, which offer larger sample sizes and greater flexibility. When used to answer appropriate research questions, database studies can also replace RCTs entirely, as database studies often result in similar findings with statistically insignificant differences. Additionally, compared to database studies, RCTs can underestimate or overestimate drug benefit/risk ratios, which affect regulatory and clinical decisions. Database studies can also be used to strengthen hypotheses in clinical trials and identify potential to repurpose existing drugs.

Although database studies can often be used in lieu of RCTs, they cannot efficaciously answer every scientific question. Determining the ideal study design begins with asking whether baseline randomization is necessary. Randomization is required if there are confounding factors, such as lack of clinical equipoise or poor measurement of outcome risk factors prior to study start. Database studies can use selection instead of randomization to separate subjects into exposure and comparator groups. The next question for determining study type is whether primary data must be collected, which is often the case when important variables are unmeasured or incompletely measured in existing data sources. If baseline randomization and primary data are not required, researchers can leverage database study designs, such as longitudinal cohort studies that use RWD. These studies can not only complement RCT data but also verify results from previous RCTs.

RWE studies can be strengthened through the following factors:

1. **Transparency**: Study methods, data sources, codes, assumptions, and study definitions must be transparent to ensure reproducibility of results. Sharing programming code
alone is not enough, because code can be built on invalid assumptions; transparency occurs when researchers provide detail on how the intention of the study and the study protocol were translated into code. The Structured Template and Reporting Tool for RWE (STaRT-RWE) and HARmonized Protocol Template to Enhance Reproducibility (HARPER) offer tools to aid reproducibility.

2. **Fit-for-purpose data**: Many RWD sources do not directly provide actual measures of study validity and data quality (e.g., sensitivity, specificity), so identifying the right combination of measures is crucial for ensuring RWE study validity.

3. **Validity**: RWE study validity refers to the likelihood that an equivalent RCT would have produced the same conclusions. The Randomized Controlled Trials Duplicated Using Prospective Longitudinal Insurance Claims: Applying Techniques of Epidemiology (RCT-DUPLICATE) program is developing methods to evaluate RWE studies by simulating RCTs and using these results to identify factors that increase RWE study validity.

4. **Improved data source linkages and access**: Linking data sources such as Medicare claims, MDS data, EHRs, and other surveys provides many opportunities for validating RWE and identifying associations. However, access costs to some data sources, such as the National Death Index, can become prohibitively expensive when examining millions of records over multiple years. Resolving these issues provides opportunities for improved RWE.

**RWD Linkages: Possibilities for NIA Longitudinal Studies** – *Maria Glymour*

Dr. Glymour began by highlighting two key takeaways. First, NIA-funded representative cohorts (e.g., the Health and Retirement Study) are useful scaffolds from which other data sources can be linked, and additional resources (e.g., cross-walking tools) must be supported to facilitate data linkages. Second, large passively accruing data sources (e.g., EHRs, claims databases) provide opportunities for low-cost, light-touch randomized studies that can predict health outcomes and establish causal relationships.

Observational research is often premised on identifying, measuring, and modeling confounders of exposures and outcomes, which has led to the common assumption that observational research is fundamentally weaker than RCTs. However, RCTs can present many difficulties, including the challenge of recruiting representative populations and the costs and efforts involved in such studies. No dataset is perfect, and there are tradeoffs in all study designs. Accordingly, data sources can best be leveraged by bridging different data types with complementary strengths.

Dementia research is often plagued by challenges with causal inference. Expanding observational instrumental variable (IV) studies offers an alternative approach to RCTs for examining causality. IV studies control for confounding and measurement error in observational studies by examining relationships between explanatory and dependent variables. IV studies have three components: (1) a source of random assignment, (2) a first-stage data source that assesses the random assignment value and type of exposure, and (3) an outcome data source that measures the random assignment value and outcome(s). Observational IV studies can use passive randomization techniques based on stable, recorded individual features of patient data,
such as the day of the month of participants’ birthdates. Additionally, the first-stage and outcome data sources need not be the same. For example, first-stage data can come from measurements of existing cohorts, and outcome data can come from routine surveillance data such as Medicare claims or VA data.

Encouragement designs slightly modify the conventional IV study design by adding an encouragement between randomization and treatment. A recent encouragement study examining effects of an FDA-approved AD prevention drug demonstrated the capabilities of low-touch IV studies based on RWD. Study participants were individuals diagnosed with MCI within a six-month window and randomized based on whether they were born on an odd or even day of the month. Participants born on odd days were sent letters describing the available medication, while those born on even days did not receive the letter. The study used CMS or HCP data to compare use of the AD drug, and outcomes were measured by examining progression of dementia in these participants from EHR or CMS data. Data were then analyzed to examine how encouragement affected rates of drug use and to examine dementia progression in both the encouragement and control groups. Such studies represent the capabilities of linking RWD sources in large, low-touch pragmatic studies.

Using RWD for Identifying and Validating Drug Repurposing Candidates for AD/ADR
– Mark W. Albers

The biology and complexity of AD/ADR presents significant challenges for drug development, particularly for therapeutics targeting pre-symptomatic disease. The natural progression of AD/ADR spans decades, with different biomarkers (e.g., amyloidosis and tau pathology) present at different stages. Furthermore, susceptibility and resilience to AD/ADR are influenced by numerous continuously evolving biological and environmental factors.

RWD analyses offer the potential to emulate RCTs in cases where RCTs are not possible. This emulation process was used to study the effects of two antidiabetic drugs—sulfonylurea and metformin initiators—on incidence of AD. Randomization was emulated by balancing baseline covariates using inverse probability of treatment weighting (IPTW) for treatment choice. Outcomes were measured by medical coding for: (1) ICD-9/10 codes that show the age of diagnosis for MCI/dementia; (2) new prescription of acetylcholinesterase inhibitors or memantine, which is used to treat dementia in AD/ADR patients; and (3) death, which is a competing risk for dementia. These analyses demonstrated that metformin initiators decreased risk of dementia onset relative to sulfonylurea initiators in two disparate cohorts in the US and the UK. Metformin was also shown to significantly increase survival advantage relative to sulfonylureas in both cohorts.

Multiple challenges remain for using RWD to identify and validate drug candidates. Drugs more recently approved by the FDA have been taken by fewer patients and have less follow-up data. Federated learning approaches can help resolve these issues by combining EHR datasets to increase sample sizes and length of follow-up. Individual health care institutions, each with a unique IRB, can analyze their own identifiable data using harmonized methods. These analyses are transferred to a central site, which returns parameters back to institutions to help them
generate a likelihood function with the effect size and variability. These likelihood functions are sent back to the centralized site, where they can be used to generate an overall effect size.

Principles of Mendelian randomization can be used to combine EHR, genomic, and claims data to identify potential drug targets. The low sensitivity of outcome data in many EHR systems for key outcomes, such as MCI and dementia, also poses challenges. These challenges can be mitigated by using natural language processing (NLP) algorithms to examine clinical notes and through combining this examination with surrogate measures such as MRI images. Overcoming these challenges can improve the ability of RWD analyses to identify and improve treatments for AD and ADRD.

**Unification of Clinical Trial Data** – *Gabriel S. Eichler*

The pharmaceutical industry is experiencing consistently reduced returns on investment in research and development (R&D) of new drugs due to long development timelines, high development costs, and high failure rates. Many pharmaceutical companies are responding to this challenge by investing in data and analytics initiatives, partnering with artificial intelligence (AI) companies, and using pooled RWD to improve the drug discovery process. Many larger pharmaceutical companies, such as Novartis, have millions of patient data points, and harmonizing and analyzing these data can transform drug development and improve clinical trial recruitment and retention.

Novartis approached this opportunity by developing the data42 platform, which integrates data from Novartis clinical trials, preclinical studies, RWD sources, and molecular data, such as genomic and proteomic datasets. data42 includes an analytics layer for Novartis users to collaborate and perform exploratory analyses, and a knowledge layer for sharing and exchanging results and analytical approaches across cross-functional teams of researchers. data42 has ingested 1 petabyte (PB) of data from approximately 2,600 clinical trials. Novartis is further developing data42 to (1) study polygenic risk scores (PRSs) for selecting target patient groups; (2) de-risk first-in-human (FIH) studies based on preclinical data; (3) develop synthetic trial arms; and (4) conduct likelihood assessments of adverse events in clinical trials.

The development of data42 highlights multiple challenges (e.g., inconsistent trial formats, different governance processes) that can inform similar large-scale RWD infrastructure efforts. These challenges can be mitigated through standardization and risk-based anonymization. Patient consent had been collected in thousands of different documents over 20 years, demonstrating a need for a broader solution to enable use of combined data while still maintaining sufficient consent and privacy. Novartis also encountered cultural hurdles, including researcher concerns about how the data will be used. Despite these challenges, data42 showcases the capability of large-scale unified data sources to guide drug development and repurposing, thereby benefiting patients.

**Panel Discussion**

*Moderator – Nina Silverberg*
Julie Zissimopoulos

Pharmaceutical companies have few financial incentives to sponsor clinical trials for repurposed generic drugs, making secondary data valuable for examining the potential for these drugs to treat AD/ADRD. The breadth of data in Medicare claims is limited, but—because these data are available for most older Americans across different demographic groups and from entry in the program until death—these data are particularly useful for quantifying the variability of drug effects on dementia risk across sex, race, and ethnicity. Researchers can also use algorithms to identify non-white populations with likely dementia. Furthermore, since most older adults take multiple medications, Medicare data can provide insight into the differential effects of distinct drug classes, drug interactions and polypharmacy effects. Combining Medicare data with other claims data into an all-payer claims database could provide further insight into dementia risk in adults prior to when they become eligible for Medicare.

Causal analyses of medication, claims, and diagnostic data to determine dementia risk must account for potential confounding factors that lead patients to take different drug classes for the same conditions for non-clinical reasons. However, there are factors associated with drug use that are potentially unrelated to dementia risk and thus can be used to support causal analysis. These include physician prescribing practices, differential access to drugs due to differences in drug penetration into different local market, cost differences, and prior authorization differences across different types of Medicare Part D plans. Modeling techniques are useful for determining how potentially small effects on dementia risk in individuals can lead to large population-level impacts.

The ability to examine cognitive decline is currently limited in Medicare data due to the lack of a reimbursed screening program for cognitive decline. Cognitive assessments, however, are included in the Annual Wellness Visit, but information on whether an assessment was performed, and how it was performed, is not included in billing data. While claims data demonstrate only if an Annual Wellness Visit occurred, they can be bolstered through linking EHR and other data.

Dementia research also requires additional research on the financial implications of dementia, including financial loss due to impaired financial decision-making and vulnerability to scams, and lost wages and productivity of family caregivers. Integrating Social Security Administration (SSA) data can provide new insights on some of these financial implications, particularly for adult children of patients with dementia.

Financial impacts are also tied to differences in reimbursement policies, including the variation in Medicare benefits offered for beneficiaries in different Medicare Advantage plans and variation in benefits and coverage across Medicare Part D plans. These plans vary significantly in coverage and cost, requiring research on how differences in coverage influence health outcomes and quality of life. RWD can be used to assess what matters—and what does not matter—for various health and financial outcomes.

New technologies for data collection and advances in the state of knowledge about AD/ADRD provides an opportunity to rethink what and how we collect and use data. For example, we
now collect real time data through devices and sensors-and with it we can employ new ‘just in
time adaptive interventions.’ We need to consider how we turn new types of data into accurate
and complete, secure, interpretable information for all stakeholders (individuals, physicians,
researchers, policymakers), and importantly make them widely available in a usable form.

**Kenneth Langa**
The Health and Retirement Study (HRS) is an NIA-and SSA-funded longitudinal panel study that
surveys a representative sample of approximately 20,000 individuals in the United States. HRS
began collecting data in 1992 and links Medicare, SSA, geographic, and genetic data, enabling
analyses across different data types. The longitudinal nature of the HRS has enabled the study
to capture data on participants prior to the onset of dementia. HRS also includes many
individuals who are unlikely or unable to participate in traditional RCTs.

HRS investigators have not conducted encouragement designs to avoid increasing participant
burden and attrition. The Alzheimer’s Association recently funded a pilot study for the HRS to
collect MRI and amyloid imaging from a subset of participants, but the protocol was met with
caution from the University of Michigan’s social science IRB, which was unfamiliar with the
pilot’s topic area. Such logistical barriers must be considered when selecting experiments for
implementation.

**Marina Sirota**
While linking EHRs can provide many insights, Dr. Sirota emphasized the need for thinking
beyond EHRs, and discussed the value of linking genotype, molecular, cohort study and imaging
data. The Sirota laboratory is currently examining the potential for linking molecular and EMR
data to examine drug repurposing, sex differences and SDOH. This work includes using EMR
data to validate hypotheses generated with molecular data. Lastly, partnerships between
academia, government, and industry can be leveraged to help solve important research
questions.

**General Discussion**
Dr. Eichler noted the interest of pharmaceutical companies in exploring longer-term follow-ups
of clinical trial participants through tokenization technologies, which can be used to track
outcomes, efficacy, and safety in the real world for years beyond the clinical trial.

Neither Medicare claims databases nor Medicare Current Beneficiary Survey (MCBS) data can
be used in isolation to examine important financial outcomes that can often serve as an early
signal of cognitive decline. Linking credit data to claims data can help researchers answer
questions about financial outcomes, but this linkage requires careful consideration of ethics
and patient protections.
Session 4: Discussion on Infrastructure Gaps and Opportunities

Invited Comments

**CMS Perspective: Shari M. Lang & Steven Farmer**

CMS aims to create a health system that achieves equitable patient-centered care, supports innovation, addresses affordability, and partners with organizations to improve health outcomes. Because most AD/ADRD patients have one or more comorbidity, CMS broadens its purview beyond the disease state itself to also include definitive health outcomes, such as longer life, significant symptom improvement, and reduced need for burdensome tests. To track Medicare beneficiaries through inpatient and outpatient settings, CMS developed a central Data Element Laboratory to standardize data collection. CMS uses collected data in numerous ways, including in clinical standards development, coverage decisions, provider enrollment, and safety oversight.

CMS bases coverage decisions on items or services that are considered “reasonable and necessary” under the Social Security Act, which requires that items and services are safe and effective, not experimental or investigational, and appropriate for Medicare beneficiaries. CMS has used RWD analyses and pragmatic clinical trials to augment RCTs as data sources for coverage decisions. For example, CMS used administrative claims data to inform its coverage decision for leadless pacemakers.

CMS is further exploring the role of fit-for-purpose studies to inform coverage decisions, given that these studies often have larger sample sizes, longer study timeframes, and more diverse participant populations than most RCTs. While CMS does not have specific guidance for fit-for-purpose studies, it does use the following general principles:

- Study protocols should be posted/published in advance
- Core robustness checks should be pre-specified
- Study execution should be rigorous, transparent, and reproducible
- Study limitations must be clearly defined
- Results must be published in a peer-reviewed, English-language journal

**Sudhir Sivakumaran**

RWD are needed to generate meaningful RWE that increases the field’s understanding of disease pathology, treatment outcomes, and many other use cases outside the laboratory. RWD reflect how drugs interact with the real world settings in which people live. Registries must be regulatory grade to ensure that data are relevant and reliable. Critical Path for Alzheimer’s Disease (CPAD) aims to connect the dots between RWD generation and meaningful RWE, and considers all types of data when developing data infrastructure and tools for treatment development. While data quality must continually be addressed, extensive infrastructure is also needed to enable analyses of RWD. RWD are large, unstructured, and have inconsistent formatting; data preparation to address these problems can include the development of data dictionaries, common data elements, and strategies that account for...
missing data. Building RWD infrastructure is an important step to bolster data collected from RCTs, particularly for neuroscience applications.

**Penny Dacks**
This workshop took place at an exciting time, when RWE has enormous potential to shift drug research from efficacy to effectiveness, deliver insight into prevention strategies, and improve inclusivity. These advantages are clear to The Association for Frontotemporal Degeneration (AFTD), which has found that the majority of individuals who want to participate in research cannot do so due to financial and physical stress. However, although RWD are a powerful tool, researchers must be aware of their limitations. As one example, data from electronic medical records (EMRs) or claims databases may be misleading in research on the causes of dementia (such as FTD) that are frequently misdiagnosed or under-diagnosed.

Additionally, trust must be both built and earned between data collectors and the individuals providing data. As RWD infrastructure is improved and data sources are increasingly integrated, that infrastructure must account for potential risks and abuse not only by the biomedical researchers and health care providers (HCPs) for whom the data infrastructure may be built, but by governments, financial institutions, insurance companies, and others who may eventually access the data infrastructure and algorithms. For example, if certain demographic groups or genotypes are identified to be at higher risk of dementia, insurance companies may increase premiums for that group. Some technologies and RWD algorithms are in development to detect dementia at its earliest manifestation of symptoms, which would lay the groundwork for pivotal trials to test prevention therapies. At the same time, our society must consider potential abuses of that technology before it is built and deployed. As one example, would financial institutions be able to deny loans or insurance to people who do not themselves know their risk? The risks and benefits of emergent RWD must be understood through expanded discussions that include not only HCPs, researchers, advocacy groups, and persons living with dementia and their care partners, but also other industries who may have an interest in these emerging types of data and technology.

**David Atkins**
The U.S. Department of Veterans Affairs (VA) has extensive experience with RWD and has used a clinical data system for over 30 years. VA data are representative across multiple dimensions (except gender; VA beneficiaries are predominantly male). VA data can be particularly useful for studying AD/ADRD, given the large number of older VA beneficiaries. The COVID-19 pandemic created a unique environment for testing capabilities for building and leveraging VA’s RWD infrastructure, and the VA has conducted more than 200 RWD studies during the past two years. For example, the VA collaborated with the FDA Evidence Accelerator to show that, in real-world settings, remdesivir does not reduce hospital stay (which was the outcome that drove the drug’s emergency use authorization). Dr. Atkins credits these accomplishments to the VA’s development of a common dataset of COVID-relevant information.

Based on the VA’s experience, Mr. Atkins believes that RWD can be best leveraged by improving data quality both within individual health systems and across multiple health
systems. A balance must be made between the value of data quality and data quantity. Similarly, data protection needs must be balanced alongside the potential information losses that occur through data deidentification. The VA experienced challenges in sharing identifiable data with external researchers; text notes have large amounts of useful information but are challenging to deidentify. Standardizing data collection will benefit both clinical care and research.

**Russ Paulsen**

In 2010, 500,000 deaths were attributed to AD; based on population growth, the current annual AD death rate is approximately 600,000 deaths per year, which approaches the scale of COVID-19. Addressing AD requires large-scale solutions, including public-private partnerships, evidence generation networks, and enhanced clinical trial infrastructure—all of which were deployed in response to the COVID-19 pandemic. During the pandemic, an Evidence Accelerator was formed that assembled data from across industries, and large collaborations put “all hands on deck” to examine COVID-19-related problems and possible solutions. However, Mr. Paulsen questioned whether AD is treated with this same urgency and creativity, and argued that stakeholders must “think big” to create RWD solutions that address AD and minimize patient and clinician burden. AD is a COVID-sized problem and large-scale solutions are needed.

**Mark A. Supiano**

RWD collection must be informed by the providers that care for older adults, and many EHR systems do not capture measures that are most relevant to patients living with dementia or their physicians. The Age Friendly Health Systems Initiative, led by the Institute for Healthcare Improvement, promotes the “4 Ms”: what matters (i.e., health outcomes most important to that patient), medication, mentation, and mobility—all of which are specific outcomes that matter to older adults.

For AD/ADRD, capturing mentation is crucial; although work to improve compliance with and completion of annual wellness visits is important, these visits must begin incorporating standardized cognitive screening. Similarly, standardized mobility screening would significantly improve risk adjustment, care planning, and screening for older adults, and measuring gait speed could advance the science of frailty. In inpatient settings, delirium is highly associated with increased mortality risk and poor health outcomes, yet hospitals rarely systematically screen for potential delirium. Hospitals can address this problem by incorporating standardized delirium screening as part of vital signs checks during inpatient care for older adults.

**Crystal M. Glover**

To prevent the perpetuation of health disparities in aging and AD/ADRD, RWD analyses must deliberately adopt the goal of assessing and addressing these disparities. Furthermore, RWD can be used within the framework of health equity, a strengths-based lens that leverages existing coping strategies and other aspects of resilience among demographically diverse persons to inform interventions to slow the acceleration of health disparities. Facilitating health equity requires incorporating demographically diverse people, with particular consideration of
intersecting identities, throughout the research process, from community outreach and study design to data analyses and interpretation, and dissemination of study findings.

Improving RWD to analyze and address health disparities in aging and AD/ADRD also requires expanding types—and minimizing hierarchies—of methods and data. Qualitative, quantitative, and mixed-methods study designs must be properly utilized to address a research purpose. Qualitative data can inform the field of social cognitive processes and cultural norms and behaviors that shape varying outcomes in aging and AD/ADRD – including uptake of repurposed drugs, medication adherence, and positive or healthy behaviors – within various demographically diverse communities that can be replicated at scale to benefit all.

Invited Commenter Discussion: Building Equitable Real-World Data Infrastructure that Addresses Wide-Ranging Stakeholder Needs

Moderators – Partha Bhattacharyya and Lisa Onken, NIA

Data Governance and Common Data Elements

Data cannot be separated from the patients providing that data, and infrastructure development should include patient perspectives. For example, which outcomes are most important to patients, and how does that inform infrastructure design? Incorporating broad and diverse patient feedback into infrastructure design can strengthen the infrastructure’s ability to meet its goals and guide the selection of data elements for inclusion. For example, studies (not yet published) from Us Against Alzheimer’s learned that patients across the AD spectrum list neuropsychiatric symptoms as being more important than memory or ADL (activities of daily living) outcomes. Selected data elements should cover the disease spectrum, and reflect disease presentation in real-world settings across representative populations.

Building an equitable RWD infrastructure first requires developing data governance models that proactively address data security, consent, privacy, and ethical concerns. Infrastructures should also determine clear and intentional goals, and only include data that are in scope of those goals. RWD infrastructure development should take into account guidance from the recently published framework from the Food and Drug Administration.

Defining common data elements is another crucial step when developing RWD infrastructure. For example, CMS has developed a Data Element Library that defines common data elements, including consensus measures for cognition and outcomes related to function. Some of these data elements efficiently aim to collect data for multiple purposes, including payment. Many of the measures in this Data Element Library are subjective, and panelists highlighted the need to complement these measures with objective assessments of key aspects such as cognition, physical performance, and function. Subjective and participant perspective data on cultural factors and the SDOH could be expanded and leveraged to examine research related to health equity.

While the Department of Health and Human Services (HHS) is working on data quality, the development of a consistent standard across the HHS that defines quality, missingness, acceptable methods, etc. would be valuable to inform common data elements. CMS does not
yet have guidance in this space, which leaves investigators making assumptions in place of following specified standards.

Adding more objective measures to patient screening risks increasing the burden on providers and their staff, which jeopardizes adherence to measurement procedures. While adding measures would increase physician/staff burden, a mini-cognitive assessment, which only takes a few minutes, could be added to an annual wellness visit, and does not require high technical knowledge to administer.

**Striving for Greater Diversity in Clinical Trials**
Incorporating and accounting for demographic diversity must occur throughout the research process, starting at the study design phase. During COVID-19, pharmaceutical companies were pushed to run diverse clinical trials, and these efforts largely succeeded. Increasing diversity in AD/ADRD trials requires a similar push from regulatory agencies. Incorporating diversity into COVID-19 clinical trials was guided by regulatory agency guidance and defined goals, and AD/ADRD research can use a similar approach to ensure diverse representation in RCTs, pragmatic clinical trials, and RWD analyses. Incorporating greater amounts of data can enable researchers to examine a diverse population with more statistical power. Additionally, the accuracy of available data can vary based on race, ethnicity, and socioeconomic status. Many disadvantaged groups unable to access specialized AD/ADRD care or treatment leave behind a less accurate and less complete medical record.

**Data Deidentification**
RWD collection, integration, and analyses need to account for the challenges presented by data deidentification, and the Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirements for deidentification, which differ depending on the research question or study design. Deidentification requires balancing the need for sufficient research data while considering privacy concerns and regulations. Based on the VA’s experience, externally sharing data is generally challenging. For some research, the VA has invited researchers to access deidentified VA data within VA systems rather than trying to export these records to other groups.

**Closing Panel: Prioritizing Gaps and Opportunities for Using RWD in AD/ADRD Research**
*Moderators* – Patricia Jones and Nadezda Radoja
*Panelists* – Malaz Boustani, David Dore, Kenneth Langa, Marina Sirota, and Julie Zissimopoulos

Panelists were asked to identify key gaps and opportunities for establishing RWD infrastructure to support AD/ADRD research.

**Marina Sirota**
Capturing relevant clinical outcomes in a standardized format and enabling interoperability between datasets remains a challenge. The COVID-19 pandemic provided lessons learned and infrastructures for data sharing and interoperability that can be applied to AD/ADRD research.
For example, drug repurposing research for COVID-19 was aided by Cerner’s ability to pool and clean data. However, data integration remains a challenge, particularly between data types (e.g., EHRs, molecular data, imaging).

Multiple institutions, such as Vanderbilt University and Geisinger Health, are developing methods to integrate molecular and EHR data on the same individual. Expanding these efforts offers the potential for more personalized medicine, such as stratifying patients by APOE genotype. Dr. Sirota’s laboratory is conducting research using public domain molecular data and EMR data to address the same research question. For example, in the case of drug-repurposing, hypotheses generated through molecular data were validated by methods using EMR data. Further integration of molecular and EHR/EMR data can identify additional insights.

In addition to identifying and studying therapeutics, RWD also provide researchers with opportunities to examine patient heterogeneity, which can lead towards more personalized treatment approaches. Applying unsupervised and clustering approaches to clinical data can offer additional opportunities. ML, predictive algorithms, and other similar methods can benefit research on AD/ADRD, though researchers must be wary of biases within various ML-based approaches.

**Malaz Boustani**
Expanding and modernizing health information exchanges for AD/ADRD offers opportunities for greater RWD sharing and integration, particularly for EMRs, claims data, death index, and other vital statistics. Furthermore, these health information exchanges should be complemented with both patient- and caregiver-reported outcomes as well as molecular data. While health information exchanges were originally established in the 1990s to improve health care, the next generation of exchanges can be designed to optimize both care and research, including support for RCTs and pragmatic trials.

More research is needed on data transportability and data generation processes, as many biases can be introduced at the data collection stage. When building data networks, advances in Agile frameworks, network science, and behavioral economics can be leveraged to enhance recruitment.

**David Dore**
The restrictions deidentification places on researchers remain a challenge, but can be addressed through two approaches. The first approach is to conduct more detailed research within HIPAA covered entities such as the UC Health system, which can provide more detailed data than the fully deidentified forms required under HIPAA. The second approach is to have individual researchers obtain approval to access data beyond what can be included in a standard deidentified dataset. Both of these approaches require appropriate human subject protections oversight, however.

Additionally, rigid standardization and exact representativeness of a sample are not necessarily needed to produce generalizable results. While a data source may not have the same proportions of different subgroups as the target population, analyses can account for these
differences as long as these subgroups are sufficiently present in the data source. Similarly, using an overly rigid approach to data standardization, particularly for demographic data from underrepresented groups, can actually further underrepresent those groups. Understanding the provenance and limitations of the data, along with knowledge representation approaches, can be used to enhance interoperability and generalizability without requiring precisely standard data definitions.

**Kenneth Langa**
Dr. Langa agreed with Dr. Boustani’s comments on promoting enhanced health information exchanges, and emphasized the importance of developing standard operating procedures for collating data from EHRs, claims, and other data sources. Additionally, data access could be improved by streamlining regulatory burden that currently makes accessibility cumbersome for researchers. More research is also needed to examine the accuracy and potential biases in the underlying data.

Encouragement trials and ongoing NIA-funded epidemiological cohort studies (e.g., Health and Retirement Study and the National Health and Aging Trends Study) offer additional opportunities for better understanding the impact of AD/ADRD treatments.

**Julie Zissimopoulos**
One additional priority is the need to identify new and less burdensome data collection methods, with the consideration that data must be transformed into useable information for researchers, physicians, patients, caregivers, and policymakers. Dr. Zissimopolous also stated that RWE must be rigorous enough to justify changing clinical care (e.g., repurposing drugs, reducing patient risk) or to inform policy around long-term care services and reimbursement.
## Appendix 1. Workshop Participants

### Presenters

**Mark W. Albers**, Assistant Professor of Neurology, Harvard Medical School; Wilkens Endowed Scholar, Department of Neurology, Massachusetts General Hospital

**David Atkins**, Director of Health Services Research & Development, Veterans Affairs

**Niranjan Bose**, Managing Director, Health & Life Sciences Strategy, Gates Ventures

**Malaz Boustani**, Professor of Aging Research, Indiana University School of Medicine

**Atul Butte**, Distinguished Professor and Director of the Bakar Computational Health Sciences Institute, University of California San Francisco; Chief Data Scientist, University of California Health System

**Maria Carrillo**, Chief Science Officer, Alzheimer’s Association

**Penny Dacks**, Senior Director of Scientific Initiatives, The Association for Frontotemporal Degeneration (AFTD)

**David Dore**, Principal Scientist, Exponent; Adjunct Professor, Brown University School of Public Health

**Gabriel S. Eichler**, VP & Head of Data, data42, Novartis

**Barbara Entwisle**, Professor of Sociology, University of North Carolina

**Crystal M. Glover**, Assistant Professor of Psychiatry and Behavioral Sciences, Rush Alzheimer’s Disease Center

**Maria Glymour**, Professor of Epidemiology & Biostatistics, University of California San Francisco

**Steven Farmer**, Chief Strategy Officer, Coverage and Analysis Group, Centers for Medicare & Medicaid Services

**Adrian Hernandez**, Duke Health Cardiology Distinguished Professor and Executive Director, Duke Clinical Research Institute; Vice Dean, Duke University School of Medicine

**Olin Janssen**, Researcher, Maastricht University

**Kenneth Langa**, Associate Director of the HRS, Co-PI of the HCAP Network, and Professor of Internal Medicine, University of Michigan; Veterans Affairs Ann Arbor Center for Clinical Management Research

**Emily Largent**, Assistant Professor of Medical Ethics and Health Policy with a Second Appointment at Penn Law, University of Pennsylvania

**Shari M. Ling**, Deputy Chief Medical Officer, Centers for Medicare & Medicaid Services

**Stephanie Monroe**, Executive Director of African Americans Against Alzheimer’s, Us Against Alzheimer’s

**Vincent Mor**, Professor of Health Services, Policy & Practice and PI of the IMPACT Collaboratory, Brown University

**Corrie Painter**, Deputy Director, Count Me In, Broad Institute

**Russ Paulsen**, Chief Operating Officer, Us Against Alzheimer’s

**Rema Raman**, Professor of Neurology and Director of Biostatistics and Recruitment at the Alzheimer’s Therapeutic Research Institution, Keck School of Medicine, University of Southern California

**Sebastian Schneeweiss**, Professor of Medicine and Epidemiology, Harvard Medical School; Chief, Division of Pharmacoepidemiology, Brigham & Women’s Hospital
Marina Sirota, Associate Professor of Pediatrics, University of California San Francisco
Sudhir Sivakumaran, Vice President of the Neuroscience Program and Executive Director of the CPAD Consortium, C-Path
Mark A. Supiano, Chief, Division of Geriatrics, University of Utah School of Medicine; Executive Director, University of Utah Center on Aging; Board of Directors, American Geriatrics Society
Michelle E. Tarver, Deputy Director of the Office of Strategic Partnerships and Technology Innovation and Program Director for Patient Science and Digital Health Center of Excellence, Center for Devices and Radiologic Health, Food and Drug Administration
Julie Zissimopoulos, Associate Professor, PI for the Center for Advancing Sociodemographic and Economy Study of ADRD and the USC RCMAR, University of Southern California

NIA RWD Working Group (*Planning Committee Members)
Office of the Director
   Richard Hodes, Director
   Melinda Kelley, Acting Deputy Director
   Patricia Jones, Director, Office of Special Populations*
Division of Aging Biology
   Ronald Kohanski, Division Director
   Stacy Carrington-Lawrence, Deputy Director*
Division of Behavioral and Social Research
   Lisbeth Nielsen, Division Director
   Partha Bhattacharyya, Chief Data Officer, BSR, NIA*
   Jonathan King, Senior Scientific Advisor to the Division Director
   Lisa Onken, Director, Behavior Change and Intervention Program*
   John Phillips, Chief, Population and Social Processes Branch
Division of Extramural Affairs
   Kenneth Santora, Division Director
   Holly Massett, Chief, Office of Clinical Research
   Joshua Park, Scientific Review Officer
   Saroj Regmi, Program Officer
Division of Geriatrics and Clinical Gerontology
   Evan Hadley, Division Director
   Sergei Romashkan, Chief, Clinical Trials Branch
   Marcel Salive, Health Science Administrator*
Division of Neuroscience
   Eliezer Masliah, Division Director
   Suzana Petanceska, Director, Office of Strategic Development and Partnerships
   Nadezda Radoja, Senior Adviser for Regulatory Science*
   Laurie Ryan, Chief, Clinical Interventions and Diagnostics Branch
   Nina Silverberg, Director, Alzheimer’s Disease Center Program*
Appendix 2. Agenda

Gaps and Opportunities for Real-World Data Infrastructure

Stakeholder Workshop
May 4, 2022

The National Institute on Aging (NIA) is convening an exploratory workshop to learn what gaps exist in current real-world data infrastructure, and what opportunities lie in expanding availability of real-world data sources for aging and Alzheimer’s Disease and Alzheimer’s Disease and Related Dementia (AD/ADRD) research. Specifically, researchers will highlight gaps and opportunities to:

1. Securely access health data (e.g., EMR, claims, genetics, etc.) from private data providers and academic institutions to gain insight on AD/ADRD disease trajectory
2. Identify opportunities for developing a platform for an AD/ADRD digital cohort that serves as a diverse recruitment pool (i.e., Recruitment as a Service) for clinical trials across NIA extramural divisions (e.g., drug trials, prevention trials, dementia care interventions)
3. Analyze sensitive RWD through secure cloud workspaces while protecting privacy of the study participants
4. Collaborate in partnerships with health care and community health providers to enable rapid drug trials (RDT) to launch embedded pragmatic clinical trials (RePCT) for improving care for older adults with multimorbidity, including dementia

In attendance will be stakeholders from the National Institutes of Health, private industry, academia, and NIA funded investigators who have first-hand knowledge of real-world data and its application. Presenters and discussants are charged with identifying and prioritizing gaps and opportunities that exist in current real-world data infrastructure that require further exploration to support aging and AD/ADRD research.

Introduction:

10:30 AM  Welcome Remarks and Charge
Richard Hodes
Director, National Institute on Aging

Session 1:  Overview of Current Infrastructure

10:40 AM  Session Plenary Speaker
Gaps and Opportunities in Current Real-World Data Infrastructure
Atul Butte
Distinguished Professor  
Director, Bakar Computational Health Sciences Institute  
University of California, San Francisco  
Chief Data Scientist, University of California Health System

11:00 AM **Real World Infrastructure, the Alzheimer’s Association National Registry, and Opportunities for Improved Patient Outcomes**  
Maria Carrillo  
Chief Science Officer, Alzheimer’s Association

11:15 AM **Real-World Evidence and Alzheimer’s Disease: The ROADMAP Data Cube**  
Olin Janssen  
Researcher, Maastricht University

11:30 AM **Key Barriers to Clinical Trials for Alzheimer's Disease**  
Niranjan Bose  
Managing Director, Health & Life Sciences Strategy, Gates Ventures

11:45 AM **Panel Discussion**  
Moderator: Jennie Larkin, NIA

David Dore, Principal Scientist, Exponent  
Barbara Entwisle, Professor of Sociology, University of North Carolina  
Stephanie Monroe, Executive Director, African Americans Against Alzheimer’s

12:05 PM **BREAK 15-MINUTES**

**Session 2  Recruitment and Clinical Trial Uses of Real-World Data**

12:20 PM **Session Plenary Speaker**  
**Overview of Possibilities of Real-World Data for Clinical Trials and Patient Experience**  
Adrian F. Hernandez  
Duke Health Cardiology Distinguished Professor  
Executive Director, Duke Clinical Research Institute  
Vice Dean, Duke University School of Medicine

12:40 PM **Race and Ethnicity in Real-World Data Sources**  
Michelle E. Tarver  
Deputy Director, Office of Strategic Partnerships and Technology Innovation  
Program Director, Patient Science Digital Health Center of Excellence  
Centers for Devices and Radiological Health, Food and Drug Administration

12:55 PM **Recruitment for Alzheimer’s Disease Prevention Trials**  
Rema Raman  
Professor of Neurology, University of Southern California

1:10 PM **Recruitment for Dementia Care Trials**  
Vincent Mor  
Professor of Health Services, Policy & Practice  
PI, IMPACT Collaboratory  
Brown University
1:25 PM  **Panel Discussion**  
Moderator: Marcel Salive, NIA

Malaz Boustani, Professor of Aging Research, Indiana University  
Emily Largent, Assistant Professor, University of Pennsylvania  
Corrie Painter, Deputy Director, Count Me In, Broad Institute

1:45 PM  BREAK 30-MINUTES

**Session 3  Secondary Data Uses of Real-World Data**

2:15 PM  **Session Plenary Speaker**  
*Which Scientific Questions can be Answered with RWD?*  
Sebastian Schneeweiss  
Professor of Medicine, Harvard Medical School  
Chief, Division of Pharmacoepidemiology, Department of Medicine, Brigham & Women’s Hospital

2:35 PM  **RWD Linkages: Possibilities for NIA Longitudinal Studies**  
Maria Glymour  
Professor and Program Director, University of California San Francisco

2:50 PM  **Using RWD for Identifying and Validating Drug Repurposing Candidates for AD/ADRD**  
Mark W. Albers  
Assistant Professor of Neurology, Harvard Medical School  
Wilkens Endowed Scholar, Massachusetts General Hospital

3:05 PM  **Unification of Clinical Trial Data**  
Gabriel S. Eichler  
VP and Head of Data, data42, Novartis

3:20 PM  **Panel Discussion**  
Moderator: Nina Silverberg, NIA  
Julie Zissimopoulos, Associate Professor, University of Southern California  
Kenneth Langa, Professor of Medicine, University of Michigan  
Research Scientist, Veterans Affairs  
Marina Sirota, Associate Professor, University of California San Francisco

3:40 PM  BREAK 15-MINUTES

**Session 4  Discussion on Infrastructure Gaps and Opportunities**

3:55 PM  **Invited Comments**  
Shari M. Ling, Deputy Chief Medical Officer, Centers for Medicare & Medicaid Services  
Steven Farmer, Chief Strategy Officer, Centers for Medicare & Medicaid Services  
Sudhir Sivakumaran, Vice-President, Neuroscience Program, Critical Path Institute  
Penny Dacks, Senior Director of Scientific Initiatives, The Association for Frontotemporal Degeneration  
David Atkins, Director of Health Services and Research Development, Veterans Affairs  
Russ Paulsen, Chief Operating Officer, Us Against Alzheimer’s
Mark A. Supiano, Chief, Division of Geriatrics, University of Utah School of Medicine
Board of Directors, American Geriatrics Society
Crystal M. Glover, Assistant Professor of Psychiatry and Behavioral Sciences, Rush
Alzheimer’s Disease Center

4:30 PM  Invited Commenter Discussion: Building equitable real-world data infrastructure that addresses wide-ranging stakeholder needs
Moderators: Partha Bhattacharyya & Lisa Onken, NIA

Session 5  Closing Panel: Prioritizing Gaps and Opportunities for Using RWD in AD/ADRD Research
Moderators: Patricia Jones & Nadezda Radoja, NIA

Malaz Boustani, Professor of Aging Research, Indiana University
David Dore, Principal Scientist, Exponent
Kenneth Langa, Professor of Medicine, University of Michigan
Research Scientist, Veterans Affairs
Marina Sirota, Associate Professor, University of California San Francisco
Julie Zissimopoulos, Associate Professor, University of Southern California