PREPARE Challenge Winner's Announcement

1. Please provide your preferred information for use in announcing the winners of the competition:

    i. Dr. Gagan Wig
        a. Name (first and last name): **Gagan Wig**
        b. Hometown: **Vancouver, Canada**
        c. A recent picture of yourself or digital avatar (feel free to attach separately):

        d. Social handle or URL (optional):
        https://x.com/GaganWig
        https://www.wigneurolab.org

    ii. Dr. Micaela Chan
        a. Name (first and last name): **Micaela Chan**
        b. Hometown: **Denton, USA / Hong Kong**
        c. A recent picture of yourself or digital avatar (feel free to attach separately):

        d. Social handle or URL (optional): https://micaelachan.com/

iii. Ziwei Zhang
    a. Name (first and last name): **Ziwei Zhang**
    b. Hometown: **Beijing, China**
    c. A recent picture of yourself or digital avatar (feel free to attach separately):

    d. Social handle or URL (optional): https://www.linkedin.com/in/ziweizhang2405/

2. Who are you (mini-bio) and what do you do professionally?

Team mini-bio:

We are a team of neuroscientists working together in a Cognitive and Systems Neuroscience laboratory based at the Center for Vital Longevity at the University of Texas at Dallas.

Dr. Gagan Wig is the principal investigator and director of the laboratory, Dr. Micaela Chan is a staff scientist in the laboratory, and Ms. Ziwei Zhang is a doctoral student completing her dissertation in the laboratory under Dr. Wig's supervision.

Our lab uses non-invasive brain imaging technologies to study brain structure and function in both human and non-human individuals. We study the organization and function of brain networks as they change over the lifespan, in both health and disease. We want to understand the biological and environmental factors that make some individuals resilient to age-related brain decline, and those that make others vulnerable to age-accompanied disease (particularly, Alzheimer's Disease). Identifying early brain and behavioral markers of risk and resilience are central to this effort, as it could lead to the development of interventions to slow or even arrest disease onset, and potentially increase the health span of aging adult individuals.

Or, individual detailed bios (if preferred):

i. Gagan is an Associate Professor of Psychology in the Center for Vital Longevity at the University of Texas at Dallas, and in the Department of Psychiatry at The University of Texas Southwestern Medical Center. Gagan earned a Bachelor of Science from the

University of British Columbia and a Ph.D. in Cognitive Neuroscience from Dartmouth College, followed by post-doctoral fellowships at Harvard University and Washington University in St. Louis with the Human Connectome Project. Gagan leads a laboratory that uses brain imaging to study healthy and pathological aging across the adult lifespan. This work is centered around measuring patterns of brain connectivity to understand how large-scale brain networks change as individuals age and determining how these changes impact cognitive function and dementia risk. Through this work the team is isolating the factors that promote resilience to age-related cognitive decline and examining how an individual's social and economic conditions can expose vulnerabilities in their brain health as they grow older.

ii. Micaela is a research scientist in the Wig Neuroimaging Lab. She completed her Ph.D. in the Cognition and Neuroscience program at UT Dallas, and her B.S. in Psychology from University of Illinois at Urbana-Champaign. Her research focuses on how our experience and environment interacts with the brain's functional network organization, and developing methods to represent individualized brain networks in a shared multivariate space. Her broader scientific interest includes network analysis, optimizing workflow and data-processing, and contributing to open-source projects/open science.

iii. Ziwei Zhang is a Ph.D. student in the cognitive neuroscience program at UT Dallas, working under the supervision of Dr. Wig. Before joining the Wig Neuroimaging Lab, she completed a bachelor's degree in psychology and worked as a research assistant in a neuroimaging lab at UCLA. Her primary research interests involve using neuroimaging methods to study the mechanisms of neurodegenerative disease. Ziwei's dissertation project examines the effects of aging and Alzheimer's disease on functional brain network organization and cognition.

3. What motivated you to compete in this challenge?

Alzheimer's Disease (AD) is devastating the lives of individuals and their loved ones. World-wide population demographics foreshadow immense individual and societal burdens of the disease. Developing tools for Alzheimer's Disease diagnosis, prediction, and intervention necessitates a collaborative effort from experts in multiple disciplines, not limited to neuroscientists and physicians.

There has considerable recent progress in understanding the brain changes that precede and characterize AD. However, it is clear that there exist brain and behavioral signals of AD that have yet to be uncovered, which would help the effort. Our research team's expertise is in examining how large-scale functional brain networks change over the lifespan and how they are impacted by AD. The Alzheimer's Disease Neuroimaging Initiative has been a tremendous resource that has enabled us and others to conduct much of this work, as the initiative has collected functional brain scans relevant to measuring an individual's brain network organization and clinical measurements on thousands of individuals varying in health status. However, incorporating these functional brain scans in predictive models requires careful processing, data cleaning, and data curation that adheres to known principles of neurobiological organization. We have been developing and applying these tools in our own work and are excited to

facilitate and catalyze others to work in this space. We recognize that providing this data will be particularly helpful for research groups with expertise in predictive modelling, but who have less familiarity with the steps needed to access, process, and shape the brain network data into a form that increases the likelihood of uncovering informative patterns of brain connectivity.

4. High level summary of your dataset: the data source, target, predictors, sample size and use for early, inclusive prediction of AD/ADRD.

   The dataset included is a processed and curated functional MRI dataset that quantifies the organization of functional brain network organization in individuals ranging in age and dementia severity. Recent work has demonstrated that the brain network patterns differ in relation to age and Alzheimer's Disease (AD) severity, and that certain changes in functional brain network organization among healthy individuals are prognostic of AD.

   The dataset is compiled from data shared by the Alzheimer's Disease Neuroimaging Initiative (ADNI). We have processed raw neuroimaging files and aggregated these outputs with clinical ratings relevant to AD dementia. The primary predictors are functional brain network relationships between 502 brain regions and the primary target variables are clinical dementia ratings and sum-of-boxes scores, obtained from clinical assessments. We also include several types of demographic information for each individual.

   The dataset includes participants with both cross-sectional data (i.e., a single scan session of data) and those with longitudinal data (i.e., multiple scan sessions collected over time), the latter of which allows for predictive models of future outcomes in individuals (e.g., individuals who are initially cognitively healthy, but were diagnosed with cognitive impairment years later). Cross-sectionally, the dataset includes 393 healthy participants (clinical dementia rating [CDR] = 0), 326 very mild dementia (CDR = 0.5), 71 mild dementia (CDR = 1) and 17 moderate dementia (CDR = 2) patients (age range: 55 – 96 years old, Female = 422); this totals to 807 unique participants. Out of the 807 participants, 385 have longitudinal MRI scans and with CDR scores at multiple timepoints (ranging from 2 to 10 scans, age range: 55 – 94 years old, Female=191).

5. What are two or three unique strengths of this dataset or type of data for early, inclusive prediction of AD/ADRD?

   i. We have provided an extensively processed resting-state fMRI dataset along with clinically relevant target variables of AD dementia severity. The utility of incorporating measures of brain network function in novel models of AD prediction necessitates careful processing and assembly of the relevant data types. Functional MRI data is notoriously susceptible to artifacts from the scanner, cardiac and respiratory signals, and head-motion, which are known to alter the correlation of resting-state signals. Our team has helped develop processing pipelines for effectively cleaning this data. The data we provided underwent preprocessing and quality checking to ensure that several known sources of artifact are minimized. In addition, we have shaped the data into a format that

adheres to neurobiological organization, which will facilitate incorporating the data into predictive models.

ii. Both cross-sectional and longitudinal data are available thus making this dataset well suited to characterization of AD dementia severity and longitudinal prediction and prognosis of dementia.

iii. ADNI data are not limited to functional MRI scans and clinical scores but also include neuroimaging scans quantifying brain structure, AD-related pathology, and metabolism. In addition, ADNI includes a wealth of other indicators based on measurement of multiple biospecimens which provide information including and not limited to genetics and proteomics. As such, there is an opportunity to enrich the models of brain function with other measures relevant to AD.

iv. ADNI is an ongoing effort, and its most recent wave of data collection is focused on enrolling individuals who are under-represented in studies of AD (particularly, racial and ethnic minorities, individuals with lower educational attainment). Thus, insight gained from the dataset we have provided can ultimately be applied, tested, and updated in additional populations to ensure generalizability of findings.

6. Did you use any tools or resources for developing your submission (e.g., to find a dataset, or explore the contents of a public dataset)?

   The dataset was accessed after our team received approval from ADNI. An in-house processing pipeline was used to pre-process the raw neuroimaging data into analyzable large-scale functional brain network matrices (see Data Processing Summary document).

7. Were there any data types or sources that you explored but didn't fit for this challenge?

   While other brain imaging modalities are also available on ADNI, many of which are already processed and shared by other groups, large scale functional brain networks provide a feature-rich dataset well suited to machine learning applications.

8. How would you improve or enrich this dataset if you had access to a big research team and an unlimited budget?

   i. **Ensuring population representativeness/generalizability:**
   There are several clear ways to improve the current dataset to increase population representativeness. First, once the dataset is available, we would process and incorporate data that will be available in ADNI4. This effort of data collection is focused on enrolling individuals who are racial and ethnic minorities and individuals with lower educational attainment. Second, relevant data is available in multiple data initiatives that are based in other countries and regions that span a wide range of demographics, which could be synthesized with the work we have conducted. This would not only increase the global population representativeness of any efforts focused on studying large-scale brain networks in relation to AD, but would also ensure that we are understanding brain network alterations in individuals who are at greater risk of AD.

ii. **Incorporating measures of function into existing models of AD:**
Measures of brain function are currently not incorporated in the dominant models of AD, but contain unique sources of variance. We would like to directly incorporate and harmonize the data we have curated with other available data types to facilitate testing of multi-modal models of AD using a single framework (e.g., measures of brain pathology [amyloid, tau], and neurodegeneration). This effort will also likely enhance AD prediction given that the various brain measures contain distinct sources of information that are informative towards AD.

iii. **Evaluation of AD-related brain network changes in clinical settings**:
There are currently no assessment tools that allow for measurement of 'brain health', akin to indicators of other biological systems (e.g., blood pressure measurement for cardiovascular health). Our work on brain network organization indicates that we have identified a relevant signal that varies in relation to both 'healthy' aging and AD. We have primarily relied on legacy datasets to uncover these patterns. However, evaluating and understanding the use of measures of functional brain networks in clinical settings necessitates further development. deployment, and testing of acquisition and analysis protocols; this would involve optimizing data collection protocols and ensuring they are collectible at scale and in multiple settings. This effort would allow us to generate normative population distributions of measures of network organization and their deviations in relation to AD and other health and environmental indicators (across a range of age that includes young- and middle-age adulthood). This enhanced effort would also allow us to understand whether the brain network changes we have isolated can be inferred and/or enhanced from other available signals, which may be lower cost and easier to collect (e.g., monitoring technologies that quantify sleep, activity, communication, behavior, etc.). We have been harmonizing available datasets that allow us to piece together relevant information, but the next phase of this work would greatly benefit from a more targeted and coordinated effort towards the stated goals.