Editorial:
Designing a Protocol Adopting an Artificial Intelligence (AI)-driven Approach for Early Diagnosis of Late Onset Alzheimer's Disease

Victor O.K. Li[1]*#, Jacqueline CK Lam[1]*#, Yang Han[1], Lawrence YL Cheung[2], Jocelyn Downey[1], Tushar Kaistha[1], Illana Gozes[3]

[1]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, China
[2]Department of Linguistics & Modern Languages, The Chinese University of Hong Kong, Hong Kong, China
[3]Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Adams Super Center for Brain Studies and Sagol School of Neuroscience, Tel Aviv University, Israel

*Corresponding authors. Address: Department of Electrical and Electronic Engineering, The University of Hong Kong, Pok Fu Lam, Hong Kong. Email: vli@eee.hku.hk; jcklam@eee.hku.hk

#The first two authors have equal contributions.

1. Literature Review

Dementia is a progressive and devastating disease characterized by the gradual loss of cognitive function. Dementia affects an individual's quality of life and increases a society's burden of care. The risk of developing dementia increases with age. As of 2020, 41.27M people across the world suffer from AD (9.5M in China and 0.16M in Israel). The latest prevalence study found that about 10% of older adults in Hong Kong suffered from dementia (Lam et al., 2008), predicting that the number of local populations suffering from dementia may triple from 100,000 in 2009 to over 300,000 in 2039 (Yu et al., 2012). Neurodegeneration is the main outcome of dementia, of which Alzheimer's Disease (AD) is the most prevalent form and a leading cause of mortality across the elderly. Unfortunately, the underlying pathogenesis and mechanism of most neurodegenerative diseases remain poorly understood. To date, diagnosis of neurodegenerative diseases prior to clinical manifestation remains challenging. While a number of cognitive, behavioral and neuroimaging parameters have been found to predict neuro-deterioration (Wong, et al., 2013; Chan et al., 2011; Cheng, et al., 2018), they are not specific to AD.

Existing AD neuroscience studies have predominantly focused on AD pathways that lead to Tau phosphorylation and Amyloid-β (Aβ) accumulation, and taken these as the prominent biomarkers of AD. However, clinical assessment tools such as Amyloid imaging or Cerebrospinal Fluid (CSF) biomarkers, remain highly expensive and invasive (McKhann et al., 2011). Increasingly, attention has been given to non-invasive testing, such as analysis of the individual's behavioral and linguistic ability, or his/her own eye gaze, for early AD or mild cognitive impairment (MCI) detection (Pan et al., 2021).

AD subjects have been seen to exhibit linguistic and acoustic deficiencies during the course of disease progression. Common deficiencies include difficulties in recalling names and using verbs, making of grammatical errors, difficulties in staying with one theme, and increasing frequencies of vocal breaks and self-corrections (Fraser et al., 2016). Four key linguistic factors, including semantic impairment, acoustic abnormality, syntactic impairment and information

impairment, were used to distinguish AD subjects from normal controls (Fraser et al., 2016). Orimaye et al. (2017) studied 99 AD subjects and 99 normal controls, using transcripts obtained from a picture description task. AD subjects were found to use a lower number of syntactic components and a higher number of lexical components (Orimaye et al., 2017). There is also a significant difference in the usage of a contiguous sequence of a number of items from a given sample of text or speech (n-grams) between the AD subjects and normal controls (Orimaye et al., 2017). The application of machine learning directed to learning cognitive-linguistic statistical patterns were found to be able to distinguish AD subjects from normal controls at 80-93% accuracy (Orimaye et al., 2017).

Behavioral studies have also reported that AD subjects exhibit impairments in gaze and ocular movements. MCI subjects were reported to have smaller pupils and poorer performance when the tasks assigned were demanding and exceeded the subjects' cognitive abilities (Nam et al., 2020). Nam et al. (2020) studied facial and eye movements between 17 early AD subjects and 17 aged-control subjects while watching a 22-minute video. AD subjects were found to have a tendency to move their faces and eyes simultaneously in a vertical direction when compared to normal subjects (Nam et al., 2020). Noiret et al. (2018) studied saccadic eye movements (SEM) among 20 AD subjects and 35 healthy older subjects by assigning them with cognitive prosaccade (PS), antisaccade (AS) and predictive saccades (PreS) tasks. AD subjects showed higher latency and higher frequencies of incorrect AS tasks and spent longer to correct their AS (Noiret et al., 2018). For PreS tasks, AD subjects showed higher gait and gait variability than normal subjects when they made anticipated saccades (Noiret et al., 2018). Close relationships were found between SEM and MMSE, and episodic memory (Noiret et al., 2018).

Existing clinical observational studies on late onset ADs (LOAD), however, are restricted in sample size, and in exploring the genetic risks of LOAD among the Han Chinese population. It remains unknown whether conclusive upstream genetic markers, and concurrent linguistic, cognitive, psychological and behavioral impairments can be identified among the MCI subjects who subsequently develop LOAD.

There are multiple international datasets, e.g., Alzheimer's Disease Neuroimaging Initiative, that provides valuable genetic and clinical data on AD, MCI subjects and healthy controls. Artificial intelligence (AI) can be used in these large datasets to efficiently search for Definitive Genetic Markers (DGMs) and concurrent cognitive, linguistic, psychological, behavioral markers for early diagnosis and treatment of AD.

2. Study Objectives

First, we aim to develop an AI-driven causal model to efficiently and accurately identify upstream DGMs of LOAD, based on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (ADNI, n.d.-a), which covers comprehensive genetic and biological data of CN, MCI and AD subjects. Second, we verify if the MCI group of the Han Chinese population in Hong Kong carrying the DGMs will have a higher probability of subsequently developing LOAD, as compared to the CN group of the same population, via a five-year longitudinal observational study. Third, we identify the cognitive, linguistic, psychological and behavioral markers concurrent to the 50 MCI subjects that carry DGMs, using 50 CN subjects as the reference, via the same observational study.

Our research hypotheses are listed as follows:

Hypothesis 1. We hypothesize that DGMs characterizing MCI subjects of the Caucasian ethnicity who later develop LOAD can be identified from comprehensive, international, open AD datasets, such as ADNI.

Hypothesis 2. We hypothesize that such DGMs found in Hypothesis 1 can also be identified in the corresponding MCI subjects of the Chinese Han ethnicity, with a high probability of developing LOAD.

Hypothesis 3. We hypothesize that concurrent cognitive, linguistic, psychological and behavioral markers can be found among the MCI subjects in the Chinese Han ethnicity who possess the DGMs.

3. Study Design

We will conduct a longitudinal, prospective, and case-control observational study. The primary outcomes of our study include (1) DGMs (upstream somatic mutations along the AD pathology) for Caucasian MCI subjects who later develop LOAD and (2) DGMs for Han Chinese MCI subjects who later develop LOAD. The secondary outcomes of our study include concurrent cognitive, linguistic, psychological, and behavioral markers to the DGMs of the MCI subjects in the Chinese Han population, who later develop LOAD.

We seek to recruit 50 MCI subjects with DGMs (identified by the study team based on Hypothesis 1) and 50 cognitively normal (CN) subjects without DGMs. Participants will be selected from the Han Chinese ethnic group in Hong Kong, following the inclusion and exclusion criteria (see Sections 3.4.1 and 3.4.2), to be recruited through the site Principal Investigator (PI), based in a public hospital in Hong Kong, with the help of the hospital-affiliated community pods and elderly centers. All screening tests, including the blood tests, will be carefully administered by professional registered nurses supervised by site PI in the hospital-affiliated pods, and all follow-up tests, including (a) Clinical and Psychological, and Cognitive and (b) Linguistic, and Behavioral, will be administered to any invited subject across two separate days by professionally trained research staff, overseen by the on-site investigators of the study.

If the participating subject has developed any early symptoms of MCI, for instance, memory loss, it is recommended that an individual (next of kin or close family member such as spouse), a "Study Partner", will (1) accompany the participant to the study visits, (2) answer questions about the participant's memory and daily functioning, and (3) assist the participant with logging in to an app for smartphone-based assessments.

Studies will be conducted only after obtaining informed consents from both the subjects and their Study Partners. The potential benefits and risks that associate with this study will be clearly laid out and explained to the subjects before obtaining their informed consents. After obtaining the informed consents, each participant will be invited to the hospital for a screening study visit, followed by a baseline study visit and a yearly follow-up study visit conducted in two separate days, depending on the results of the screening blood test.

For the screening study visit, each participant will be asked to (1) provide information concerning his/her demographics, family history, medical history, and medications, (2) have a general physical examination to check his/her weight, height, and vital signs, (3) have a neurological examination, (4) provide up to 20 ml of blood drawn from his/her upper limb by an experienced research nurse. The blood samples will also be used for screening tests,

including haematology, chemistry panel, vitamin B12, and thyroid-stimulating hormone. Only CN and MCI subjects who have passed the screening tests will be recruited to undergo a genetic test to identify whether they carry DGMs. Whole exome sequencing of a screened subject's de-identified blood sample will be conducted via an HKU-affiliated genetic test centre.

After screening, a baseline study visit will be conducted for all enrolled subjects, including those who are CN (without DGMs) and those who are MCI (with DGMs), in the first year of the study. Follow-up studies will be conducted for CN (every other year) and MCI (every year). For the baseline study visit and the follow-up study visits, each subject will be asked to (1) provide his/her medication information, (2) have a physical examination (weight, height, and vital signs), (3) have an electrocardiogram (ECG) and ocular examination, (4) have a cognitive-linguistic assessment (assisted by audio and video recording), (5) have a psychological assessment (assisted by audio and video recording), and (6) have a behavioral assessment (assisted by audio and video recording). Further, a follow-up blood and genetic test will be conducted for both CN and MCI.

The screening and baseline study visits will collect initial results based on the blood and genetic test, and the validated cognitive-linguistic, psychological, and behavioural tests (see Appendices 1-18). These cognitive-linguistic, psychological, and behavioural tests assess one's memory, thinking, emotional well-being, functioning, and behaviour and have already been validated either by (1) the ADNI study, or (2) Hong Kong-based units administering neurodegeneration assessment. During the follow-up study visits, any changes in the initial results will be measured.

To reduce the length of time the subjects spent on clinical or non-clinical tests, we will ensure that each study visit consists of four rounds on two separate days (each round last for 3 hours only and can be stopped at any point at the request of the subject or study partner). Two rounds will be conducted on the same day, each last for only 3-hours followed by a 2-hour break, with lunch provided to the subjects and their Study Partners to ensure that they are given enough time for rest.

In addition to the on-site study visits, each enrolled subject will be invited to use a smartphone app developed by us to perform tasks and respond to surveys between the study visits (on a voluntary basis). To ensure personal privacy and security, the app will perform active and passive data collection tasks for cognitive-linguistic and behavioral assessments (Kourtis et al, 2019). To ensure personal privacy and security, we will adopt a trust-based privacy preserving method to protect our subjects (Crowcroft & Gascón, 2018). Permissions for both active and passive data collection will be sought. NO data access will be conducted without the subject and his/her Study Partner's prior online approval.

3.1 Sample Size
200 subjects will be recruited for screening. 100 subjects will be enrolled in the study after screening. The sample size calculation is based on the following rationales:

We plan to recruit 50 MCI subjects (with DGMs identified by the team based on the ADNI dataset) and 50 CN subjects (without DGMs). This sample size calculation is based on the findings reported by previous literature. First, using data mining techniques, Ivashko-Pachima et al. (2019) was able to identify a novel activity-dependent neuroprotective protein (ADNP) hotspot mutation from RNA and DNA samples obtained from 20 AD and 20 CN subjects. Second, using machine learning models and feature importance analysis, Chen et al. (2019)

collected smartphone-based data from 113 subjects (MCI/AD: 31; CN: 82) in the US and identified a group of behavioral and language-related features that differentiate symptomatic subjects from healthy subjects. Further, an on-going smartphone/wearable-based AD study in the EU suggested that a sample size of 55 for each study group yields at least 80% power to detect a difference in daily behaviours between the healthy and the symptomatic subjects (RADAR-AD, n.d.). Given that the older individuals in Hong Kong normally show much higher compliance, we expect more longitudinal information to be collected and analyzed for cognitive-linguistic and behavioral signatures in our study. Therefore, we determine that 100 subjects are sufficient to curate a dataset that can characterize the genetic, cognitive-linguistic, and behavioural markers of the MCI subjects in Hong Kong.

In order to obtain 100 subjects enrolled for our study, 200 subjects will be recruited for screening. This sample size is calculated based on the statistics shown in the ADNI dataset, where 70% of the AD subjects, 50% of the MCI subjects, and 42% of CN subjects have APOE ε4-negative. APOE ε4 is a widely known genetic risk factor of AD. Assuming our DGMs occur at a similar rate, say in 50% of the subjects, we expect to be able to find 50 subjects with DGMs after sampling 100 subjects, and 50 without the DGMs after sampling another 100 subjects.

3.2 On-site Assessment

Blood Sampling and Genetic Profiling. We will use the blood samples from each of MCI subjects to identify 50 subjects with identified genetic markers, and CN subjects to identify 50 normal controls without genetic markers. The blood samples will be collected as in Malishkevich et al. (2016) and subjected to RNA and DNA preparation followed by RNA-seq (Ivashko-Pachima et al., 2019), as well as DNA whole exome sequencing for verification.

Cognitive-Linguistic-Psychological Assessment. The following assessment tools were used in the ADNI3 study protocol (ADNI, n.d.-b). Based on these tools, our study will develop a cognitive-psychological-linguistic assessment for the Hong Kong context. The questionaries (see Appendices 1-17) include Cognitive Change Index (CCI), Mini Mental State Examination (MMSE), Cogstate Brief Battery (CBB), Logical Memory I and II, Everyday Cognition (ECog), Montreal Cognitive Assessment (MoCA), Category Fluency (Animals), Trails A & B, Boston Naming Test (BNT), Multi-Lingual Naming Test (MlNT), Auditory Verbal Learning Test (AVLT), Geriatric Depression Scale (GDS-15), Clock drawing, Neuropsychiatric Inventory (NPI), ADAS-Cog 13 (w/Delayed Recall & Number Cancellation), Clinical Dementia Rating Scale (CDR), and Financial Capacity Instrument – Short Form (FCI-SF).

Behavioral Assessment. Our study will assess personal behaviors using the speech and behavioral information (such as gait, voice, and head/eye movements) obtained from the audio/video recording. In addition, the subject will perform the Activities of Daily Living (FAQ) test used in the ADNI3 study protocol (ADNI, n.d.-b).

Following the ADNI3 study protocol (ADNI, n.d.-b), the details of the on-site study timeline are listed in Tables 1 and 2 below:

Table 1. On-site study visits for CN participants

Name	Screening	Baseline	Year 1	Year 2	Year 3	Year 4
Explain study and obtain consent	X	X				

Assessment	1	2	3	4	5
Demographics & family history	X				
Medical history, physical & neurological examination, Hachinski Ischemic Score	X				
Weight & vital signs	X	X		X	X
Height	X				
Screening labs (hematology, chemistry panel, Vitamin B_{12}, TSH)	X				
DNA and RNA sample collection	X			X	X
ECG examination		X		X	X
Ocular examination		X		X	X
Concomitant medications	X	X		X	X
Adverse events	X	X		X	X
Diagnostic summary	X			X	X
Cognitive-Linguistic-Psychological Assessment					
Clinical Dementia Rating (CDR)	X			X	X
Mini Mental State Examination (MMSE)	X			X	X
Montreal Cognitive Assessment (MoCA)		X		X	X
Neuropsychiatric Inventory Score (NPI)		X		X	X
Everyday Cognition (ECog)		X		X	X
ADAS-Cog 13 (w/Delayed Recall and Number Cancellation)		X		X	X
Cogstate Brief Battery (CBB)		X		X	X
Logical Memory I and II	X			X	X
Cognitive Change Index (CCI)		X		X	X
Clock drawing		X		X	X
Trails A & B		X		X	X
Boston Naming Test		X		X	X
Category Fluency (Animals)		X		X	X
Auditory Verbal Learning Test		X		X	X
Multi-Lingual Naming Test (MlNT)		X		X	X
Financial Capacity Instrument – Short Form (FCI-SF)		X		X	X
Geriatric Depression Scale (GDS-15)	X			X	X
Behavioral Assessment					

Circadian rhythms, gait, voice, and eye and head movements		X		X		X
Activities of Daily Living (FAQ)		X		X		X

Table 2. On-site study visits for MCI participants

Name	Screening	Baseline	Year 1	Year 2	Year 3	Year 4
Explain study and obtain consent	X	X				
Demographics & family history	X					
Medical history, physical & neurological examination, Hachinski Ischemic Score	X					
Weight & vital signs	X	X	X	X	X	X
Height	X					
Screening labs (hematology, chemistry panel, Vitamin B_{12}, TSH)	X					
DNA and RNA sample collection	X		X	X	X	X
Electrocardiogram (ECG) examination		X	X	X	X	X
Ocular examination		X	X	X	X	X
Concomitant medications	X	X	X	X	X	X
Adverse events	X	X	X	X	X	X
Diagnostic summary	X		X	X	X	X
Cognitive- Linguistic-Psychological Assessment						
Clinical Dementia Rating (CDR)	X		X	X	X	X
Mini Mental State Examination (MMSE)	X		X	X	X	X
Montreal Cognitive Assessment (MoCA)		X	X	X	X	X
Neuropsychiatric Inventory Score (NPI)		X	X	X	X	X
Everyday Cognition (ECog)		X	X	X	X	X
ADAS-Cog 13 (w/ Delayed Recall and Number Cancellation)		X	X	X	X	X
Cogstate Brief Battery (CBB)		X	X	X	X	X
Logical Memory I and II	X		X	X	X	X
Cognitive Change Index (CCI)		X	X	X	X	X
Clock drawing		X	X	X	X	X

Trails A & B		X	X	X	X	X
Boston Naming Test		X	X	X	X	X
Category Fluency (Animals)		X	X	X	X	X
Auditory Verbal Learning Test		X	X	X	X	X
Multi-Lingual Naming Test (MlNT)		X	X	X	X	X
Financial Capacity Instrument – Short Form (FCI-SF)		X	X	X	X	X
Geriatric Depression Scale (GDS-15)	X		X	X	X	X
Behavioral Assessment						
Circadian rhythms, gait, voice, and eye and head movements		X	X	X	X	X
Activities of Daily Living (FAQ)		X	X	X	X	X

3.3 Smartphone/Wearable-based Assessment

Cognitive-Linguistic-Psychological Assessment. Smartphone-based cognitive-linguistic assessment will be conducted on a voluntary basis. Every three months, each subject will be invited to fill in cognitive-linguistic questionnaires and perform speaking tests via his/her smartphone microphone. They will be asked to perform tests, such as describing a picture or remembering words after listening to them. The audio recording will only happen when the subjects are conducting language-related tests. An audio recording icon will flash in the app as the recording is ongoing. The app will not record any voice information in the background.

Behavioral Assessment. Wearable-based behavioral assessment will be conducted on voluntary basis. Each subject will be provided with a wearable device such as FitBit and be invited to put on his/her wrist to collect personal behavior and activity data (such as circadian rhythms and gait) continuously for one week per month. This passive data collection procedure aligns closely with the cardiovascular health research app developed by Stanford University (McConnell et al., 2017). The app, named MyHeart Counts, has been validated by Stanford University's research team for activity data collection in the United States, United Kingdom, and Hong Kong and has sought online privacy and security approval before use. We will follow the same procedure to ensure that online passive data access is on par with the strictest international privacy and security laws and regulations.

3.4 Selection and Withdrawal of Subjects
3.4.1 Subject Inclusion Criteria
The following inclusion criteria for both CN and MCI are based on the ADNI3 study protocol (ADNI, n.d.-b) with additional criteria being included to capture the Hong Kong context.

3.4.1.1 CN Subject Inclusion Criteria
- Clinician concluded that the patient is cognitively normal in a key public hospital in HK.
- Normal memory function documented by scoring above education adjusted cutoffs on the Logical Memory II subscale (Delayed Paragraph Recall, Paragraph A only) (the maximum score is 25):

a. ≥9 for 16 or more years of education
b. ≥5 for 8-15 years of education
c. ≥3 for 0-7 years of education

- Mini-Mental State Exam (MMSE) score between 23 and 30 (inclusive). (Exceptions may be made for subjects with less than 2 years of education. For instance, the cut-off should be >18 for subjects without education and > 20 for those with education but less than 2 years).
- Clinical Dementia Rating = 0. Memory Box score must be 0.
- Cognitively normal, based on an absence of significant impairment in cognitive functions or activities of daily living.
- Stability of Permitted Medications[1] for 4 weeks:
  a. Stable doses of antidepressants lacking significant anticholinergic side effects (if they are currently adequately treated for depressive symptoms and do not have a history of major depression within the past 1 years)
  b. Estrogen replacement therapy is permissible
  c. Gingko biloba is permissible, but discouraged
  d. Washout from psychoactive medication (e.g., excluded antidepressants, neuroleptics, chronic anxiolytics or sedative hypnotics, etc.) for at least 4 weeks prior to screening

3.4.1.2 MCI Subject Inclusion Criteria

- Subject must have a subjective memory concern as reported by subject, study partner, or clinician.
- Abnormal memory function documented by scoring below education adjusted cutoffs on the Logical Memory II subscale (Delayed Paragraph Recall, Paragraph A only) (the maximum score is 25):
  a. < 11 for 16 or more years of education
  b. ≤ 9 for 8-15 years of education
  c. ≤ 6 for 0-7 years of education
- Mini-Mental State Exam (MMSE) score between 23 and 30 (inclusive). (Exceptions may be made for subjects with less than 2 years of education. For instance, the cut-off should be >18 for subjects without education and > 20 for those with education but less than 2 years).
- Clinical Dementia Rating = 0.5. Memory Box score must be at least 0.5.
- General cognition and functional performance sufficiently preserved such that a diagnosis of AD cannot be made by the site physician at the time of the screening visit.
- Stability of Permitted Medications[2] for 4 weeks:
  a. Stable doses of antidepressants lacking significant anticholinergic side effects (if they are currently adequately treated for depressive symptoms and do not have a history of major depression within the past 1 years)
  b. Estrogen replacement therapy is permissible
  c. Gingko biloba is permissible, but discouraged
  d. Washout from psychoactive medication (e.g., excluded antidepressants, neuroleptics, chronic anxiolytics or sedative hypnotics, etc.) for at least 4 weeks prior to screening
  e. Cholinesterase inhibitors and memantine are allowable if stable for 12 weeks prior to Screening Visit

---

[1] This refers to the subject's own medications, not prescribed by our research team.
[2] This refers to the subject's own medications, not prescribed by our research team.

3.4.1.3 Additional Subject Inclusion Criteria
- Chinese Han ethnicity
- Geriatric Depression Scale less than 6.
- Age between 55-90 (inclusive).
- Study partner is available who has frequent contact with the subject (e.g. an average of 10 hours per week or more) and can accompany the subject to all clinic visits for the duration of the protocol.
- Visual and auditory acuity adequate for neuropsychological testing.
- Good general health with no diseases expected to interfere with the study.
- Participant is not pregnant, lactating, or of childbearing potential (i.e. women must be two years post-menopausal or surgically sterile).
- Hachinski Ischemic Score less than or equal to 4.
- Must speak Cantonese fluently.

3.4.2 Subject Exclusion Criteria
The following exclusion criteria for both CN and MCI subjects are adapted from the ADNI3 study protocol (ADNI, n.d.-b).
- For CN, any significant neurologic disease, such as Parkinson's disease.
- For MCI, any significant neurologic disease other than suspected incipient Alzheimer's disease, such as Parkinson's disease.
- Participants with multiple lacunes or lacunes in a critical memory structure.
- Major depression, bipolar disorder as described in DSM-IV within the past one year. Psychotic features, agitation or behavioral problems within the last three months which could lead to difficulty complying with the protocol.
- Currently treated with medication for obsessive-compulsive disorder or attention deficit disorder.
- History of schizophrenia (DSM IV criteria).
- History of alcohol or substance abuse or dependence within the past two years (DSM IV criteria).
- Any significant systemic illness or unstable medical condition which could lead to difficulty complying with the protocol.
- Residence in skilled nursing facility.
- Current use of specific psychoactive medications (e.g., certain antidepressants, neuroleptics, chronic anxiolytics or sedative hypnotics).
- Current use of any other exclusionary medications.
- Participation in clinical studies involving neuropsychological measures being collected more than one time per year.

3.4.3 Subject Withdrawal Criteria
The subjects can withdraw from the study at any time without giving a reason. The site investigators will make every reasonable effort to maximize participant retention. Subjects who discontinue early from the study will be encouraged to have an early termination visit at the point of discontinuation. The early discontinuation visit will contain the same assessments as the complete annual visit, to allow collection of the main outcome measures.

4. Data Analysis
4.1 Hypothesis 1
To examine our Hypothesis 1, using existing AD datasets, we will identify the DGMs via an AI-driven causal graph model developed by us. Our model consists of three components that

cover data collection, deep learning model development, and AI-based causal graph development. Firstly, we collect data from the most comprehensive longitudinal and publicly accessible ADNI dataset that carries whole genome sequencing, demographics, and cognitive scores of 810 subjects with highly variable trajectories of cognitive change, over an eight-year period. Secondly, we develop a hybrid deep learning (HDL) model, trained on the ADNI (and other datasets) and labelled with CN, MCI, and AD to obtain the best parameters to fit the data. Thirdly, we develop a Bayesian deep causal graph to further identify the most causal factors (those with the highest probability) of these high-potential biological/genetic markers as DGMs of LOAD. The highest-potential genetic markers, demographics (e.g., age and gender), cognitive-linguistic-psychological features, and other most probable confounders (e.g., medical history) will serve as the nodes of a causal graph. The causal relations are denoted as the directed edges. Potential causal graphs with the highest probabilities will then be generated using causal graph construction techniques. Graphs aligned with our expert biomedical knowledge will be identified for further analysis. Using our selected AD datasets, we will apply de-confounding techniques to remove confounders and determine the DGM nodes with causal links to LOAD (Pearl, 2009). Such DGMs will be verified by our longitudinal observational study in Part II.

4.2 Hypothesis 2
To examine our Hypothesis 2, based the data collected in our longitudinal observational study, statistical analysis, ANOVA or linear regression model will be used to determine whether there is a statistically significant difference ($p<0.05$) in terms of the onset of LOAD for our MCI group vs. our CN group, properly adjusted in terms of age, education, cognitive abilities, such as MMSE score, etc., at the end of the five-year period.

4.3 Hypothesis 3
To examine our Hypothesis 3, based the data collected in our longitudinal observational study, an AI model will be developed, trained with input data consisting of linguistic, cognitive, psychological, and behavioral markers, labelled as having DGMs, or without DGMs, properly adjusted with confounding factors such as age, education, cognitive ability (e.g., MMSE score). Based on the saliency scores of the different linguistic, cognitive, psychological, and behavioral markers, we can determine which combination of markers are concurrent to DGMs. Thus, such markers can be used as early detectors of LOAD.

5. Data Processing and Storage
5.1 Data Consent, Security, and Privacy
All information collected during the course of the study visits will be kept strictly confidential. The subject's personal identifiers (name and address) will be replaced by a unique research identifier, which is a randomly generated alphanumeric code that is unique to each research participant in this study. The name and other personal information which identify the subject will not be linked to his/her research data. The raw coded data (from surveys and tests) will be computerized and transmitted to and stored in a secure cloud server with strict security measures. Moreover, a strong authentication system is in place to secure access to the encrypted data.

Any information collected from the smartphone app will be kept strictly confidential and will be the minimum required by the study. For video data collected from the cognitive-psychological-linguistic and behavioral assessments, we will adopt a number of privacy protection techniques, such as body masking, face blurring, and cartooning, for de-identification, while preserving the key information (such as eye/head movements) (Climent-

Perez et al., 2020; Haque et al., 2020). In the unlikely event of data leakage, no personal information will be associated with the data, and no individual participants can be re-identified from the data. The coded data collected from his/her smartphone will also be transferred to and stored in the same secure cloud server.

Meanwhile, blood samples will first be collected and stored securely in a refrigerator in the hospital. The refrigerator which serves as blood bank is secured by a door lock and taken care of by clinically trained research staff. Only de-identified blood sample will be sent to a specialized genetic testing unit for whole exome sequencing and destroyed immediately after the completion of genetic tests.

5.2 Pre-existing Publicly Accessible Data from Human Subjects

Our study will utilize a pre-existing dataset, namely, the ADNI dataset (ADNI, n.d.-a). The dataset can be obtained following an approval process. Consistent with the purpose of the original ADNI study, we will use the ADNI dataset to identify definitive biomarkers for LOAD. The ADNI dataset includes basic demographics and medical history, brain images, genetics, cognitive tests, cerebrospinal fluid and blood biomarkers, covering AD, MCI, and normal control subjects. The dataset is de-identified, i.e., no personal identifiers, such as name and address, are available in the data. No new data will be collected from the subjects in the original ADNI study.

To identify upstream genetic markers, other pre-existing publicly available genetic datasets, such as ADNI-Japan, ADNI-China, and signaling pathway datasets, such as KEGG (https://www.genome.jp/kegg/pathway.html), etc., will be deployed.

We may also make use of other publicly accessible datasets. All datasets will be deployed only after going through an internationally recognized approval process.

6. Study Limitations and Recommendations

Several study limitations and ways of addressing those limitations are discussed below as follows:

Representation of the majority of Hong Kong population might be inadequate. To ensure proper representation, we will extend our search via existing datasets that represents ethnicities closer to the Han Chinese, such as the ADNI-Japan dataset.

Further, relying on ADNI dataset alone to identify the right DGMs might not be exhaustive. We will boost our findings by using existing, ever growing datasets other than ADNI to complement the existing ADNI dataset to extend our search for the right DGMs.

Missing data are expected from concurrent cognitive-linguistic-psychological and behavioural markers data collection. A novel probabilistic classification method can be developed to fill in such missing data (Li, Han, Li, and Lam, et al., 2021). Other relevant missing data approaches such as Long-Short Term Context Encoder and other AI-driven imputation approaches taking advantage of the domain-specific knowledge, such as the high correlation between age onset and education, can be exploited (Yu, Li, and Lam, 2021; Yu, Yu, Li, and Lam, 2020).

Study recruitment and participation may present a psychological barrier due to a paucity of medication. However, families of AD patients as well as carriers of the major genetic risk allele,

apolipoprotein E4 (APOE4), may be willing to participate in future developments of relevant prophylactic and disease modifying therapies.

In future, an AI-driven causal model, integrating genetic, clinical, linguistic, cognitive, psychological and behavioural assessments, in a longitudinal fashion, will accelerate the identification of upstream genetic markers and high propensity linguistic-cognitive-psychological markers characterizing early-stage AD, and speed up drug repurposing that lead to the identification of a combination of potential drug candidates from FDA approved drugs.

Acknowledgements

References

1. Yu, R., Chau, P. H., McGhee, S. M., Cheung, W. L., Chan, K. C., Cheung, S. H., & Woo, J. (2012). Trends in prevalence and mortality of dementia in elderly Hong Kong population: projections, disease burden, and implications for long-term care. International Journal of Alzheimer's Disease, 2012.
2. Lam, LC, Tam, CW, Lui, VW, Chan, WC, Chan, SS, Wong, S, Wong, A, Tham, MK, Ho, KS, Chan, WM, Chiu, HF. (2008). Prevalence of very mild and mild dementia in community-dwelling older Chinese people in Hong Kong. International Psychogeriatrics, 20, 135-148.
3. Wong, CH, Leung, GT, Fung, AW, Chan, WC, Lam, LC. (2013). Cognitive predictors for five-year conversion to dementia in community-dwelling Chinese older adults. International Psychogeriatrics, 25, 1125-1134.
4. Chan, WC, Lam, LC, Tam, CW, Lui, VW, Leung, GT, Lee, AT, Chan, SS, Fung, AW, Chiu, HF, Chan, WM. (2011). Neuropsychiatric symptoms are associated with increased risks of progression to dementia: a 2-year prospective study of 321 Chinese older persons with mild cognitive impairment. Age and Ageing, 40, 30-35.
5. Cheng, CP, Cheng, ST, Tam, CW, Chan, WC, Chu, WC, Lam, LC. (2018). Relationship between cortical thickness and neuropsychological performance in normal older adults and those with mild cognitive impairment. Aging and Disease, 9 (6),1020-1030.
6. McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., ... & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia, 7(3), 263-269.
7. Pan, Y., Shea, Y. F., Li, S., Chen, R., Mak, H. K. F., Chiu, P. K. C., ... & Song, Y. Q. (2021). Prevalence of mild behavioural impairment: a systematic review and meta-analysis. Psychogeriatrics, 21(1), 100-111.
8. Fraser, K. C., Meltzer, J. A., & Rudzicz, F. (2016). Linguistic features identify Alzheimer's disease in narrative speech. Journal of Alzheimer's Disease, 49(2), 407-422.
9. Orimaye, S. O., Wong, J. S., Golden, K. J., Wong, C. P., & Soyiri, I. N. (2017). Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. BMC bioinformatics, 18(1), 1-13.
10. Nam, U., Lee, K., Ko, H., Lee, J. Y., & Lee, E. C. (2020). Analyzing Facial and Eye

Movements to Screen for Alzheimer's Disease. Sensors, 20(18), 5349.

11. Noiret, N., Carvalho, N., Laurent, É., Chopard, G., Binetruy, M., Nicolier, M., ... & Vandel, P. (2018). Saccadic eye movements and attentional control in Alzheimer's disease. Archives of Clinical Neuropsychology, 33(1), 1-13.

12. The Alzheimer's Disease Neuroimaging Initiative (ADNI). (n.d.-a). Access Data and Samples. Available at: http://adni.loni.usc.edu/data-samples/access-data/

13. Kourtis, L. C., Regele, O. B., Wright, J. M., & Jones, G. B. (2019). Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. NPJ Digital Medicine, 2(1), 1-9.

14. Crowcroft, J., & Gascón, A. (2018). Analytics without tears or is there a way for data to be anonymized and yet still useful?. IEEE Internet Computing, 22(3), 58-64.

15. Malishkevich, A., Marshall, G. A., Schultz, A. P., Sperling, R. A., Aharon-Peretz, J., & Gozes, I. (2016). Blood-borne activity-dependent neuroprotective protein (ADNP) is correlated with premorbid intelligence, clinical stage, and Alzheimer's disease biomarkers. Journal of Alzheimer's Disease, 50(1), 249-260.

16. Ivashko-Pachima, Y., Hadar, A., Grigg, I., Korenková, V., Kapitansky, O., Karmon, G., ... & Gozes, I. (2019). Discovery of autism/intellectual disability somatic mutations in Alzheimer's brains: mutated ADNP cytoskeletal impairments and repair as a case study. Molecular Psychiatry, 1-15.

17. The Alzheimer's Disease Neuroimaging Initiative (ADNI). (n.d.-b). ADNI3 ProtocolVersion3.1. Available at: http://adni.loni.usc.edu/methods/documents/

18. McConnell, M. V., Shcherbina, A., Pavlovic, A., Homburger, J. R., Goldfeder, R. L., Waggot, D., ... & Ashley, E. A. (2017). Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts Cardiovascular Health Study. JAMA cardiology, 2(1), 67-76.

19. Pearl, Judea. Causality. Cambridge University Press, 2009.

20. Chen, R., Jankovic, F., Marinsek, N., Foschini, L., Kourtis, L., Signorini, A., ... & Trister, A. (2019, July). Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2145-2155).

21. RADAR-AD. (n.d.). Remote Assessment of Disease and Relapse – Alzheimer's Disease. Available at: https://cordis.europa.eu/project/id/806999

22. Climent-Perez, P., Spinsante, S., Mihailidis, A., & Florez-Revuelta, F. (2020). A review on video-based active and assisted living technologies for automated lifelogging. Expert Systems with Applications, 139, 112847.

23. Haque, A., Milstein, A., & Fei-Fei, L. (2020). Illuminating the dark spaces of healthcare with ambient intelligence. Nature, 585(7824), 193-202.

24. Li, J. Han, Y., Lam, J.C.K., Li, V.O.K., Matthews, S., Cheung, L.Y.L., Yip, V., Downey, J., Chan, D., and Gozes, I. (2021). Correlation between the Bilingual Status and the Onset Age of AD and MCI Subjects: Evidence from the ADNI dataset. Preprint.

25. Yu, Y.W., Li, V.O.K., and Lam, J.C.K. (2021). Missing air pollution data recovery based on Long-Short Term Context Encoder. IEEE Trans. on Big Data, 2021.

26. Yu, Y.W., Yu, J.Q., Li, V.O.K. and Lam, J.C.K. (2020). A novel Interpolation-SVT approach for recovering missing low-rank air quality data. IEEE Access, Vol. 8, No. 1.