

# Screening over Speech in Unselected Populations for Clinical Trials in AD (PROSPECT-AD): Study Design and Protocol

A. König<sup>1,2,3</sup>, N. Linz<sup>3</sup>, E. Baykara<sup>3</sup>, J. Tröger<sup>3</sup>, C. Ritchie<sup>4</sup>, S. Saunders<sup>4</sup>, S. Teipel<sup>5,6</sup>, S. Köhler<sup>5</sup>, G. Sánchez-Benavides<sup>7,8,9</sup>, O. Grau-Rivera<sup>7,8,9</sup>, J.D. Gispert<sup>7,8,10</sup>, S. Palmqvist<sup>11,12</sup>, P. Tideman<sup>11,12</sup>, O. Hansson<sup>11,12</sup>

1. Institut national de recherche en informatique et en automatique (INRIA), Stars Team, Sophia Antipolis, Valbonne, France; 2. Cobtek (Cognition-Behaviour-Technology) Lab, FRIS-University Côte d'azur, Nice, France; 3. ki:elements GmbH, Saarbrücken, Germany; 4. Edinburgh Dementia Prevention, University of Edinburgh, Edinburgh, UK; 5. Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Rostock, Germany; 6. Department of Psychosomatic and Psychotherapeutic Medicine, University of Rostock, Rostock, Germany; 7. Barcelonaβeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain; 8. IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain; 9. Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable (CIBERFES), Madrid, Spain; 10. Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain; 11. Clinical Memory Research Unit, Department of Clinical Sciences, Malmö, Lund University, Lund, Sweden; 12. Memory Clinic, Skåne University Hospital, Sweden

Corresponding Author: Alexandra König, ki:elements GmbH, Am Holzbrunnen 1a, D-66121 Saarbrücken, Email: alexandra.koenig@ki-elements.de

## Abstract

**BACKGROUND:** Speech impairments are an early feature of Alzheimer's disease (AD) and consequently, analysing speech performance is a promising new digital biomarker for AD screening. Future clinical AD trials on disease modifying drugs will require a shift to very early identification of individuals at risk of dementia. Hence, digital markers of language and speech may offer a method for screening of at-risk populations that are at the earliest stages of AD, eventually in combination with advanced machine learning. To this end, we developed a screening battery consisting of speech-based neurocognitive tests. The automated test performs a remote primary screening using a simple telephone.

**OBJECTIVES:** PROSPECT-AD aims to validate speech biomarkers for identification of individuals with early signs of AD and monitor their longitudinal course through access to well-phenotyped cohorts.

**DESIGN:** PROSPECT-AD leverages ongoing cohorts such as EPAD (UK), DESCRIBE and DELCODE (Germany), and BioFINDER Primary Care (Sweden) and Beta-AARC (Spain) by adding a collection of speech data over the telephone to existing longitudinal follow-ups. Participants at risk of dementia are recruited from existing parent cohorts across Europe to form an AD 'probability-spectrum', i.e., individuals with a low risk to high risk of developing AD dementia. The characterization of cognition, biomarker and risk factor (genetic and environmental) status of each research participants over time combined with audio recordings of speech samples will provide a well-phenotyped population for comparing novel speech markers with current gold standard biomarkers and cognitive scores.

**PARTICIPANTS:** N= 1000 participants aged 50 or older will be included in total, with a clinical dementia rating scale (CDR) score of 0 or 0.5. The study protocol is planned to run according to sites between 12 and 18 months.

**MEASUREMENTS:** The speech protocol includes the following neurocognitive tests which will be administered remotely: Word List [Memory Function], Verbal Fluency [Executive Functions] and spontaneous free speech [Psychological and/or behavioral symptoms]. Speech features on the linguistic and paralinguistic level will be extracted from the recordings and compared to data from CSF and blood biomarkers, neuroimaging, neuropsychological evaluations, genetic profiles, and family history. Primary candidate marker from speech will

be a combination of most significant features in comparison to biomarkers as reference measure.

Machine learning and computational techniques will be employed to identify the most significant speech biomarkers that could represent an early indicator of AD pathology. Furthermore, based on the analysis of speech performances, models will be trained to predict cognitive decline and disease progression across the AD continuum.

**CONCLUSION:** The outcome of PROSPECT-AD may support AD drug development research as well as primary or tertiary prevention of dementia by providing a validated tool using a remote approach for identifying individuals at risk of dementia and monitoring individuals over time, either in a screening context or in clinical trials.

*Key words:* Dementia, Alzheimer's disease, screening, cognitive assessment, speech biomarker, phone-based, machine learning.

## Background

Speech is a complex process involving the vocal communication of thoughts, ideas, and emotions using language by the means of spoken words and phrases. Evidence suggests that changes in language and speech characteristics may occur synchronously with progression of neurodegenerative disorders (1). These changes can cause noticeable variations in speech assessed by measures that capture what is being said through semantics, word selections and grammar usage (linguistic features) and how it is being said based on phonetic and acoustic characteristics (para-linguistic features). The term 'speech' encompasses here both of these layers. Consequently, several studies have investigated the potential use of digital biomarkers extracted from speech for early diagnosis and disease monitoring in neurodegenerative dementias (2-5).

Recent advances in establishing Brain Health Services (6), allied to the global efforts in developing disease modifying treatments for Alzheimer's disease (AD) (7), together with the scientific progress of understanding

AD pathology decades before dementia manifests (8), have resulted in a shift to identify individuals at earlier stages of the AD continuum. The importance of early detection has resulted in a need for innovative solutions that are specifically geared towards facilitating the population wide screening of at-risk individuals that are at a preclinical or prodromal stages of AD.

Technology has become an integral part of our society and most individual's everyday lives'. Commonly, participants screened for neurodegenerative disease research are exposed to a battery of costly and sometimes invasive tests to assess eligibility to join the studies or clinical trials. These tests are often administered at late disease stages and include complex cognitive assessments delivered face-to-face with a trained examiner. However, studies have shown that automated screening processes, including automated speech recognition, reaches similar accuracy as human raters (9, 10) and are also validated in a telephone-based setting (11). Furthermore, Artificial Intelligence (AI) empowered assessment pipelines can capture clinically relevant features beyond what is available in traditionally administered neurocognitive tests (12, 13). As most previous studies of speech in AD have focused on discriminating between patients who are in mild to moderate dementia stage of AD and healthy controls (14), little is known about speech changes in the preclinical phase of AD. Therefore, the current study focuses on the earliest stage of AD, a critical period in the disease processes when abnormal amyloid levels could be a more specific indicator of disease than later stages when age-related amyloid accumulation may be present.

We hypothesize that a primary screening, to detect the preclinical AD phenotype (i.e. individuals with positive AD biomarkers, but without cognitive impairment) may be possible by using speech during remotely-administered neuropsychological tests. Phone-based speech assessments would utilise advanced machine learning and aid screening in future AD trials. For larger application, a screening battery consisting of speech-based neurocognitive tests was conceived, enabling the entire first-line screening process to be performed remotely using an ordinary telephone connected to a chatbot. With recent advancements in the field of computational linguistics, more fine-grained analysis of individuals' performances may be possible (9), capturing clinically relevant features well beyond traditional manual scoring of neurocognitive tests.

The overall aim of this study is to build speech-based machine learning models for the detection of the relevant phenotype through access to multiple gold-standard phenotyped cohorts. We further aim to evaluate the validity and feasibility of the screening approach for diagnosis based on information extracted from the and participants' attitude towards its usefulness for remote screening. The exploratory aim is to investigate the free speech generated in testing sessions in regards to the presence of neuropsychiatric symptoms.

## **Primary Study Objectives**

The primary objective of this study is to develop algorithms to identify speech biomarkers for Alzheimer's disease in a population with a cognitive status ranging from cognitively unimpaired to very mild dementia.

To create an algorithm to identify speech biomarkers of Alzheimer's dementia, we will leverage data collected in previous studies (with study specific agreements in place for future data sharing) to compare speech patterns to known biomarkers of Alzheimer's disease, such as cerebrospinal fluid (CSF) amyloid- $\beta$  1-42 and CSF p-Tau. If speech biomarkers can be identified in preclinical and prodromal AD or AD dementia, the likelihood increases of selecting only the most at-risk patients and research participants for intensive and invasive gold standard biomarker methods, such as lumbar punctures and PET scans.

## **Secondary study objectives**

The secondary aims of this study are to confirm the psychometric validity of the chosen tasks when performed on the phone and understand participant's experiences of this testing medium.

Whilst many cognitive tests are validated for face to face assessments, few are validated for assessment via the phone. It is important to understand how these tasks perform in this novel environment and what ceiling effects are seen. This is important to understand as the ability to employ a fully automated system both reduces specialist staff costs and guarantees identical administration of the task for each participant.

The potential benefit of automated telephone-based testing has both immediate (research testing able to take place regardless of lockdown guidelines in place in response to the Covid-19 pandemic) and long term (involvement of participants who previously wouldn't have been able to engage with research projects if they couldn't access a site due to health or geographical limitations) implications.

## **Exploratory Objectives**

The exploratory objective of the study is to investigate whether changes in speech and language can be detected in the free speech generated in testing sessions (narrative storytelling task) in regards to the potential presence of neuropsychiatric symptoms such as depression or apathy.

For this, questions will be asked that are capable of eliciting emotional reactions (or a lack thereof) by describing events that triggered recent affective arousal.

## **Methods**

The multi-centre PROSPECT-AD study is a prospective longitudinal observational study. It will evaluate the use

of automatic speech analysis for cognitive assessments over the telephone in a study population currently included in ongoing observational cohort studies, encompassing regular in person cognitive testing. The assessments will consist of widely applied and validated neurocognitive tasks - verbal fluency and list learning - that are based on speech. Additionally, there will be free speech recorded for exploratory analysis. Participants will be asked to complete multiple phone assessments during the study period during which audio will be recorded for later analysis.

## Study design

PROSPECT-AD collaborates with already ongoing cohorts such as EPAD (UK), DESCRIBE and DELCODE (Germany), BioFINDER Primary Care (Sweden) and Beta-AARC (Spain) by adding the collection of speech data to existing protocols or as follow-up assessments over the telephone. Participants at preclinical stages are mainly recruited from existing parent cohorts across Europe to form an international 'probability-spectrum' population covering the entire continuum of anticipated probability for Alzheimer's dementia development. This characterization of cognitive, biomarker and risk factor (genetic and environmental) status of each research participants over time combined with audio recordings of speech samples will provide the necessary well-phenotyped population for developing predictive longitudinal models for AD covering the entire disease course and concurrently create a pool of highly characterized individuals for the validation analysis. The study protocol is planned to run over 18 months.

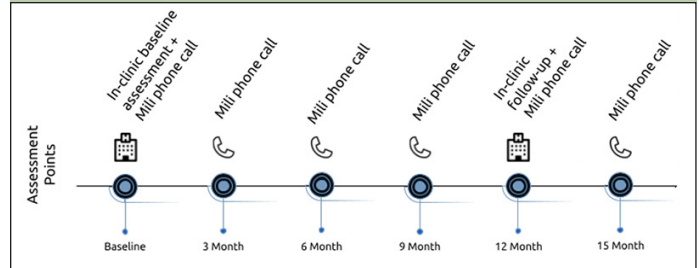
## Study protocol

At the baseline (months 0 (M0)) and follow up phone calls (between 3 to 6 months after baseline call according to cohort, +/- 14 days) participants will be called from the software (see Figure 1 for an overview of the study protocol). At the start of this automated assessment the chatbot will confirm that the participant consents to continue with the phone call. Once consent has been confirmed after providing extensive information about the study, the participant will be reminded that the phone call will be audio recorded and that they should not share any identifiable information during the call. After this the audio recording will be started. The computer will then explain the tasks to the participant, ensuring the participant is aware of the challenges they may face completing the tasks and encouraging them to try their best at the tasks. The phone assessment includes the following tasks:

- Verbal learning encoding (immediate)
- Semantic verbal fluency
- Verbal learning recall (delayed)
- Narrative storytelling task

Each of these tasks will have verbatim instructions, which will be read word by word to the participant before the task starts. The participant will be asked if they are clear on what they need to do. Afterwards, the task can start. Each of these tasks will be recorded in a secondary audio stream which just records the participant responses to allow for deep speech analysis of performance on these tasks. When all tasks have been completed the participant will be thanked for their time and efforts.

**Figure 1.** Study protocol of PROSPECT-AD. Automated speech assessments (Mili phone call) were added to the already existing study protocols



## Study population

PROSPECT-AD will recruit 300 participants from the following cohorts: EPAD (UK), DESCRIBE and DELCODE respectively (Germany), and BioFINDER Primary Care (Sweden) and 100 participants from Beta-AARC (Spain) (1000 participants in total) aged 50 or older, who have been referred to a memory clinic for evaluation of cognitive problems. In each cohort, the status of participants on the AD continuum is confirmed by CSF or PET amyloid biomarkers and CDR scores. Accordingly, every participant is selected and evaluated based on both pathophysiological (amyloid) and cognitive function axes.

## Inclusion criteria

- Age  $\geq$  50 years
- Cognition: cognitively unimpaired to MCI (CDR score max. 0.5) (2)
- Presence of diagnosis-specific biomarkers
- Fluent in the given language
- Ability to give consent or have consenting caregiver(s)
- Signed and dated informed consent form

## Exclusion criteria

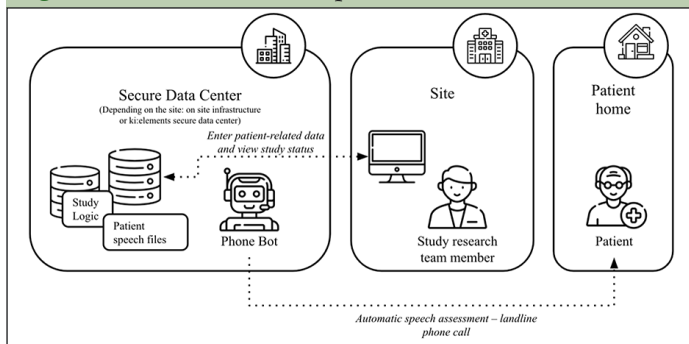
- Significantly impaired hearing ability
- Significant unstable systemic illness or organ failure that makes it difficult to participate
- Current significant alcohol or substance misuse
- Refusing investigation at the Memory clinic
- The cognitive impairment can with certainty be explained by another condition or disease such as significant anemia, infection, severe sleep deprivation,



psychotic disorder, moderate-severe depression, alcohol abuse etc.

which is done in-clinic face to face, can be found in Table 2.

**Figure 2.** Overview of the speech data collection flow



### Data collection

Speech data will be collected at each assessment time point. All data will be collected via the Mili platform over the phone. There will be no paper source documentation for this study.

In the platform, the fully automated system uses software (“Mili” phone) to connects with participants who receive an ordinary phone call to complete cognitive assessments. The software guides the interaction through a predefined protocol. The protocol includes short cognitive tests (e.g., episodic memory word list test).

The platform software used for the study is developed by ki:elements and its implementation follows high security standards and common regulatory requirements so that they are ready to use even for clinical trials (Computer System Validation, GCP guidelines, FAIR principles, GDPR requirements).

The details of data collection at Month 0 and Month 12,

### Data Management

All Investigators and study site staff involved with this study must comply with the requirements of the appropriate data protection legislation (including the General Data Protection Regulation and Data Protection Act) with regard to the collection, storage, processing and disclosure of personal information.

### Personal Data

The following data about participants will be collected:

- Name (at the local study center)
- Phone Number (at the local study center and ki:elements software)
- Age (ki:elements software)
- Gender (ki:elements software)
- Education in years (ki:elements software)
- Native and other languages (ki:elements software)
- Audio recording (ki:elements software)

Participants will be provided with information regarding the study in the participant information sheet and asked to explicitly consent to the storage of this data in these databases.

The phone calls to the participants will be made through the ki:elements software. The data is recorded and then transferred to a safe server designated for the study center (EPAD - Edinburgh server, DESCRIBE/ DELCODE – DZNE Rostock server and BioFINDER Primary Care – ki:elements server hosted by Google

**Table 1.** Study Protocol for Prospect AD Studies

Procedure	Baseline	Month 3	Month 6	Month 9	Month 12	Month 15
Informed consent	✓					
Demographic data	✓					
<b>Clinical Assessment</b>						
Medical history	✓				✓	
Biomaterial collection	✓				✓	
Brief physical & neurological examination	✓				✓	
Clinical Dementia Rating	✓				✓	
Cohort specific additional assessments (see Table 2)	✓				✓	
<b>Cognitive Assessment</b>						
Cohort specific cognitive assessments (see Table 2)	✓				✓	
<b>Speech Assessment</b>						
Fully-automated Mili Phone Assessment	✓	✓	✓	✓	✓	✓
• Verbal learning encoding (immediate)						
• Semantic verbal fluency						
• Verbal learning recall (delayed)						
• Narrative storytelling task (DESCRIBE, DELCODE, & BioFINDER Primary Care)						

**Table 2.** Cohort Specific Assessments at Month 0 and Month 12

Cohort	Patient population	Assessments
EPAD (UK)	Cognitively unimpaired to MCI	<p><b>Clinical and cognitive data</b></p> <ul style="list-style-type: none"> <li>• Spontaneous speech assessment*</li> <li>• Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)*</li> <li>• TabCat tasks: <ul style="list-style-type: none"> <li>- Favourites Recall</li> <li>- Delay Recall</li> <li>- Recognition</li> <li>- Dot Counting</li> <li>- Flanker</li> </ul> </li> <li>• 4 Mountains Test</li> <li>• Virtual Reality Supermarket Trolley</li> </ul> <p><b>Self-reported Questionnaires</b></p> <ul style="list-style-type: none"> <li>• Geriatric Depression Scale (GDS)</li> <li>• State-Trait Anxiety Inventory (STAI)</li> <li>• Healthy Ageing Through Internet Counselling in the Elderly (HATICE)</li> <li>• Swedish National study on Aging and Care (SNAC)</li> <li>• The Pittsburgh Sleep Quality Index</li> <li>• Lifestyle questionnaire</li> <li>• Covid-19 questionnaire</li> <li>• Research Experience Questionnaire</li> <li>• Personality using Big Five Inventory-10 (BFI-10)</li> <li>• Brain Injury Screening Questionnaire (BISQ)</li> <li>• International Physical Activity Questionnaire (IPAQ)</li> <li>• Social Networking Questionnaire (SNQ)</li> <li>• Amsterdam Instrumental Activities of Daily Living Questionnaire (AIADLs)</li> </ul> <p><b>Biomaterial</b></p> <ul style="list-style-type: none"> <li>• SST Gold (Serum)</li> <li>• EDTA Lavender (DNA)</li> <li>• EDTA Lavender (Blood Count)</li> <li>• EDTA Lavender (Plasma/Buffy Coat)</li> </ul>
DESCRIBE (Germany) DELCODE (Germany)	Cognitively unimpaired to MCI	<p><b>Clinical and cognitive data</b></p> <ul style="list-style-type: none"> <li>• Mini-Mental State Examination (MMSE)</li> <li>• Consortium to Establish a Registry for Alzheimer’s Disease (CERAD)</li> <li>• Neuropsychological Testing (NPT) factor scores</li> <li>• Geriatric Depression Scale (GDS)</li> <li>• Geriatric Anxiety Inventory (GAI)</li> <li>• Neuropsychiatric Inventory Questionnaire (NPI-Q)</li> <li>• Alzheimer’s Disease Assessment Scale (ADAS)</li> <li>• Measurement of Everyday Cognition (ECog)</li> <li>• NPT long-term consolidation</li> </ul> <p><b>Biomaterial</b></p> <ul style="list-style-type: none"> <li>• CSF Biomarker (Abeta 40 &amp; 42 and p-Tau)</li> <li>• ApoE e4 Genotyp</li> </ul> <p><b>MRI / PET</b></p> <ul style="list-style-type: none"> <li>• Structural MRI</li> <li>• PET scans (Amyloid +/- status)</li> </ul>
BioFINDER Primary Care (Sweden)	Cognitively unimpaired to MCI	<p><b>Clinical and cognitive data</b></p> <ul style="list-style-type: none"> <li>• The Montreal Cognitive Assessment (MOCA)</li> <li>• Mini-Mental State Examination (MMSE)</li> <li>• The 10-word list from ADAS-cog (immediate, delayed and recognition recall)</li> <li>• Symbol Digit Modalities Test (SDMT)</li> <li>• Animal Fluency</li> <li>• Cube copying</li> <li>• The clock drawing test</li> <li>• Immediate and delayed recall of pictures test</li> <li>• Trail Making Test A and B</li> <li>• Mini-Cog</li> <li>• CANTAB Paired Associate Learning (iPad)</li> <li>• CANTAB Reaction Time (iPad)</li> <li>• Mezurio Gallery game (smartphone)</li> <li>• Mezurio Tilt Task (smartphone)</li> <li>• Digital Clock Test</li> <li>• Cognitive symptom questionnaire</li> <li>• Amsterdam iADL scale</li> </ul>

**Table 2 (continued).** Cohort Specific Assessments at Month 0 and Month 12

Cohort	Patient population	Assessments
		<b>Biomaterial</b> <ul style="list-style-type: none"> <li>• CSF</li> <li>• Plasma</li> </ul> <b>MRI / PET</b> <ul style="list-style-type: none"> <li>• 18F-flutemetamol PET (small subset where CSF is not available)</li> </ul>
Beta-AARC (Spain)	Cognitively unimpaired to MCI	<b>Clinical and cognitive data</b> <ul style="list-style-type: none"> <li>• Subjective cognitive decline</li> <li>• Cognitive reserve Questionnaire</li> <li>• Family history</li> <li>• Life style habits</li> <li>• Brief COPE-28</li> <li>• NEO-FFI</li> <li>• Attitudes regarding aging – ERA 12</li> <li>• Mini-Mental State Examination (MMSE)</li> <li>• Memory Alteration Test (M@T)</li> <li>• Free and Cued Selective Reminding Test (FCSRT)</li> <li>• WMS-IV-Logical Memory</li> <li>• WAIS-IV Visual Puzzles, Matrix reasoning and Coding</li> <li>• Trail Making Test (TMT)</li> <li>• 5 digit test</li> <li>• Semantic Fluency</li> <li>• Altoida digital biomarker</li> </ul> <b>Biomaterial</b> <ul style="list-style-type: none"> <li>• CSF Amyloid Status and p-tau measures</li> <li>• Blood Based Biomarkers</li> <li>• APOE e4 status</li> </ul> <b>MRI / PET</b> <ul style="list-style-type: none"> <li>• 3T MRI</li> </ul>

These are assessments that are cohort specific, and are done at month 0 and month 12 in addition to what is stated in Table 1; \* Audio was recorded during these tasks.

Germany GmbH in Frankfurt). Ki:elements saves the collected electronic data in an encoded, password-protected database. The webhosting institute will ensure adequate physical and environmental safety including storage in rooms with strictly limited physical access. The server infrastructure offers sufficient network security and safe server access points. Remote access to the server is only possible via SSH-protocol (SSL), which uses Public-Private-Key-Encryption or similar authentication services. Relevant pseudonymized clinical data will be stored, used and processed by third parties for research purposes within the scope of the project agreement. We obtained ethics approval for three cohorts to ensure study's reasonableness for participants and accordance with national and European data security laws.

### Data Information Flow

The project will collect speech data from participants as part of the study protocol using the ki:elements software. This raw data will be stored locally on the study site server and transferred to Google Germany GmbH server (Frankfurt, Germany) for analysis by ki:elements. The speech data will be converted to numerical features at which point the data is considered pseudonymised. The analyzed speech data is owned by the specific study site and used with licensing agreements by ki:elements and the Alzheimer Drug Discovery Foundation. Data from

previous studies will be transferred to ki:elements to be combined with speech data and used for algorithm development work. Participants who were previously enrolled in the specific cohort have consented to data sharing for future studies/will be asked to provide consent as part of this study for their previously acquired data to be used for this study.

### Proposed Data analyses

To achieve the primary objective of the study, three different types of variables will be computed and extracted from speech recordings:

1. Classical neuropsychological outcome variables, i.e. total score and subscores (immediate and delayed recall) of verbal memory task, number of correctly named words in fluency tasks, repetitions.

2. Novel or qualitative outcome variables will be derived from speech recordings of the cognitive tests. These are automatically processed using the proprietary speech analysis pipeline from ki:elements that involves automatic speech recognition to transcribe and to extract semantic and linguistic information. For the verbal fluency task this includes variables such as semantic clusters which are determined using predefined semantic subcategories, mean cluster size is calculated as the sum of cluster sizes divided by the number of clusters, the number of switching clusters is defined as the total

number of switches between clusters, and temporal clusters, represent temporal alignments of produced words. Further, semantic distance between produced words is calculated using neural word embedding models and word frequency which is based on frequently occurring versus unusual words is calculated from large text corpora in a given language. For the verbal learning task, measures such as repetitions, temporal and serial clustering and primacy and recency item counts will be calculated.

3. Low-level speech descriptors will be extracted separately from the audio files from different main areas: temporal features including measures of speech proportion (e.g., length of pauses and length of speaking segments), the connectivity of speech segments and general speaking rate; prosodic features relating to long-time variations and rhythm in speech (e.g., perceived pitch and intonation of speech); formant features represent the dominant components of the speech spectrum and carry information about the acoustic resonance of the vocal tract and its use; source features relate to the source of voice production, the airflow through the glottal speech production system. Spectral features characterize the speech spectrum; the frequency distribution of the speech signal at a specific time instance information in some high dimensional representation. These features operationalize irregularities in vocal fold movement (e.g., measures of voice quality).

These measures will be computed on a task-level, i.e. for each cognitive task performed. A combination of predictive statistical and machine learning models will be used, with patients' biomarker status as a target variable. Models will be constructed with task specific features and using aggregated variables spanning multiple cognitive tasks.

To achieve the secondary objective of assessing reliability, classical neuropsychological outcome variables will be extracted from the recorded audio. To this end, human raters will use classical scoring schemes to score speech recordings of cognitive tasks. This will lead to a dataset where each participant has multiple repeated measurements of the same task. Results from multiple time points will be compared using repeated measures ANOVA and repeated measures correlation.

For the explorative objective, audio features (para-linguistic - or low-level speech descriptors as described above) will automatically be extracted from the audio signal and textual features (linguistic) from transcripts of the free speech task and compared to classical standard measures of neuropsychiatric symptoms. For this, we will perform correlation analysis and train regression models to evaluate the predictive diagnostic power of the extracted markers.

Participants' attitude towards the automated phone-based assessment will be gathered through questionnaires and interviews in order to evaluate its feasibility.

## Conclusion

Screening for trial candidates typically uses speech-based cognitive tests in a face to face setting on site which can be burdensome and very stressful for patients. However, these examinations still represent the cornerstone of cognitive assessments but have the potential to be conducted remotely at large scale decentralised trials using a telephone or smart devices.

Recently speech analysis has become mature enough to automate such speech-based testing procedures and capture clinically fine-grained features that are highly sensitive to early cognitive changes (1). Hence, speech can be utilized to indicate potential trial candidates at low cost and at a population-wide scale, and to select the right candidates to advance in the screening funnel for an on-site more comprehensive assessment stage. PROSPECT-AD aims to scale up and to validate a novel, digital, and solely speech-based biomarker extracted from an ordinary telephone to enable remote screening of early AD subjects.

The outcome of PROSPECT-AD may have a major impact on the improvement of drug development research methodology by providing a validated solution for neurocognitive screening and monitoring of participants of early AD clinical trials. Through access to large-scale, well phenotyped cohorts of early-stage patients, it can be investigated whether a speech biomarker extracted from cognitive task performances over the phone is able to automatically detect the relevant phenotype for potential enrollment in novel decentralised AD trials and non-pharmacological, preventive treatments.

*Conflict of interest:* AK is partially employed by the company ki:elements. NL, EB and JT are employed by the company ki:elements. Within the last three years ST has served on national and international advisory boards for Roche, Eisai, Biogen, and Grifols. He is member of the independent data safety and monitoring board of the study ENVISION (sponsor: Biogen). SP has served on scientific advisory boards and/or given lectures in symposia sponsored by Bioartec, Biogen, Eli Lilly, Geras Solutions, and Roche. SS receives salary from Sony Research and holds a research grant from the Scottish Neurological Research Fund. OH has acquired research support (for the institution) from ADx, AVID Radiopharmaceuticals, Biogen, Eli Lilly, Eisai, Fujirebio, GE Healthcare, Pfizer, and Roche. In the past 2 years, he has received consultancy/speaker fees from AC Immune, Amylyx, Alzpath, BioArctic, Biogen, Cerveau, Fujirebio, Genentech, Novartis, Roche, and Siemens. SK has nothing to disclose.

*Funding:* This study is funded by the Alzheimer's Drug Discovery Foundation Diagnostic Accelerator program.

*Acknowledgements:* The authors would like to thank and acknowledge all cohort investigators and their team involved in PROSPECT-AD for their help in the development and harmonization of the study protocol design.

*Ethical standards:* All human procedures will be conducted in accordance with the Declaration of Helsinki and Good Clinical Practice.

## References

1. de la Fuente Garcia S, Ritchie CW, Luz S. Artificial Intelligence, Speech, and Language Processing Approaches to Monitoring Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's disease: JAD*. 2020;78(4):1547-74; DOI: 10.3233/JAD-200888.
2. Forbes-McKay K, Shanks MF, Venneri A. Profiling spontaneous speech decline in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatrica*.

- 2013;25(6):320-7; DOI: 10.1017/neu.2013.16.
3. Hoffmann I, Nemeth D, Dye CD, Pákási M, Irinyi T, Kálmán J. Temporal parameters of spontaneous speech in Alzheimer's disease. *International journal of speech-language pathology*. 2010;12(1):29-34; DOI: 10.3109/17549500903137256.
  4. Ahmed S, Haigh AM, de Jager CA, Garrard P. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain : a journal of neurology*. 2013;136(Pt 12):3727-37; DOI: 10.1093/brain/awt269..
  5. Sanborn V, Ostrand R, Ciesla J, Gunstad J. Automated assessment of speech production and prediction of MCI in older adults. *Applied neuropsychology Adult*. 2022;29(5):1250-7; DOI: 10.1080/23279095.2020.1864733.
  6. Altomare D, Molinuevo JL, Ritchie C, Ribaldi F, Carrera E, Dubois B, et al. Brain Health Services: organization, structure, and challenges for implementation. A user manual for Brain Health Services—part 1 of 6. *Alzheimer's Research & Therapy*. 2021;13(1):168; DOI: 10.1186/s13195-021-00827-2.
  7. Cummings J, Lee G, Nahed P, Kamar M, Zhong K, Fonseca J, et al. Alzheimer's disease drug development pipeline: 2022. *Alzheimer's & dementia (New York, N Y)*. 2022;8(1):e12295; DOI: 10.1002/trc.12295.
  8. Abubakar MB, Sanusi KO, Ugusman A, Mohamed W, Kamal H, Ibrahim NH, et al. Alzheimer's Disease: An Update and Insights Into Pathophysiology. *Front Aging Neurosci*. 2022;14:742408; DOI: 10.3389/fnagi.2022.742408.
  9. König A, Linz N, Tröger J, Wolters M, Alexandersson J, Robert P. Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task. *Dement Geriatr Cogn Disord*. 2018;45(3-4):198-209; DOI: 10.1159/000487852.
  10. König A, Satt A, Sorin A, Hoory R, Derreumaux A, David R, et al. Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People. *Current Alzheimer research*. 2018;15(2):120-9; DOI: 10.2174/1567205014666170829111942.
  11. Tröger J LN, König A, Robert P, Alexandersson J. Telephone-based dementia screening I: Automated semantic verbal fluency assessment. 12th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2018;New York, USA
  12. Linz N, Tröger J, Alexandersson J, König A, editors. Using neural word embeddings in the analysis of the clinical semantic verbal fluency task. IWCS 2017-12th International Conference on Computational Semantics; 2017.
  13. Linz N, Tröger J, Alexandersson J, Wolters M, König A, Robert P, editors. Predicting Dementia Screening and Staging Scores from Semantic Verbal Fluency Performance. 2017 IEEE International Conference on Data Mining Workshops (ICDMW); 2017 18-21 Nov. 2017.
  14. Abdalla M, Rudzicz E, Hirst G. Rhetorical structure and Alzheimer's disease. *Aphasiology*. 2018;32(1):41-60.

© Serdi 2023

How to cite this article: A. König, N. Linz, E. Baykara, et al. Screening over Speech in Unselected Populations for Clinical Trials in AD (PROSPECT-AD): Study Design and Protocol. *J Prev Alz Dis* 2023;2(10):314-321; <http://dx.doi.org/10.14283/jpad.2023.11>