

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Protocol for a remote data collection speech analysis study in people at risk for Alzheimer's disease dementia: the SPeAk study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-052250
Article Type:	Protocol
Date Submitted by the Author:	13-Apr-2021
Complete List of Authors:	Gregory, Sarah; The University of Edinburgh Centre for Clinical Brain Sciences Linz, Nicklas; German Research Centre for Artificial Intelligence Saarbrücken Branch König, Alexandra; National Institute for Research in Computer Science and Automation Langel, Kai; Janssen Healthcare Innovation Pullen, Hannah; The University of Edinburgh Centre for Clinical Brain Sciences Luz, Saturnino; The University of Edinburgh Usher Institute of Population Health Sciences and Informatics Harrison, John; Metis Cognition Ltd; Vrije Universiteit Amsterdam, Department of Neurology Ritchie, Craig; The University of Edinburgh Centre for Clinical Brain Sciences
Keywords:	Dementia < NEUROLOGY, Delirium & cognitive disorders < PSYCHIATRY, Old age psychiatry < PSYCHIATRY

SCHOLARONE™
Manuscripts

1
2
3 Protocol for a remote data collection speech analysis study in people at risk for
4 Alzheimer's disease dementia: the SPeAk study
5
6

7 Sarah Gregory¹, Nicklas Linz², Alexandra König³, Kai Langel⁴, Hannah Pullen¹, Saturnino
8 Luz⁵, John Harrison^{6, 7, 8}, Craig Ritchie¹
9

10 ¹ Edinburgh Dementia Prevention, Centre for Clinical Brain Sciences, University of
11 Edinburgh, Edinburgh, United Kingdom

12 ² KI Elements, Saarbrücken, Germany

13 ³ National Institute for Research in Computer Science and Automation (INRIA), France

14 ⁴ Janssen Clinical Innovation, Spain

15 ⁵ Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

16 ⁶ Metis Cognition Ltd, Kilmington Common, United Kingdom

17 ⁷ Alzheimer Center Amsterdam, Department of Neurology, Amsterdam UMC, Vrije
18 Universiteit, Amsterdam, The Netherlands

19 ⁸ Institute of Psychiatry, Psychology & Neuroscience, King's College London, United
20 Kingdom
21
22

23
24 Email addresses: Sarah.Gregory@ed.ac.uk; nicklas@ki-elements.de; alexandra@ki-
25 elements.de; KLangel1@its.jnj.com; Hannah.Pullen@ed.ac.uk; s.luz@ed.ac.uk;
26 johnnpc@btinternet.com; craig.ritchie@ed.ac.uk
27

28
29 Corresponding author: Sarah Gregory, Edinburgh Dementia Prevention, Biocube 1, 9 Little
30 France Road, Edinburgh, EH16 4UX; Sarah.Gregory@ed.ac.uk
31

32 Keywords: speech, biomarkers, Alzheimer's disease, cognition
33

34 Word count: 2586 words
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Introduction: Identifying cost effective, non-invasive biomarkers of Alzheimer's disease (AD) is a clinical and research priority. Speech data is easy to collect, and studies suggest it can identify those with AD. We do not know if speech features can predict AD biomarkers in a preclinical population.

Methods and analysis: The Speech on the Phone Assessment (SPeAk) study recruits participants aged 50 years and over who have previously completed studies with AD biomarker collection. Participants complete a baseline telephone assessment including spontaneous speech and cognitive tests. A 3-month visit will repeat the cognitive tests with a conversational artificial intelligence bot. Participants complete acceptability questionnaires after each visit. Participants are randomised to receive their cognitive test results either after each visit or only after they have completed the study. We will combine SPeAK data with AD biomarker data collected in a previous study and analyse for correlations between extracted speech features and AD biomarkers. The outcome of this analysis will inform the develop of an algorithm for prediction of AD risk based on speech features.

Ethics and dissemination: This study has been approved by the Edinburgh Medical School Research Ethics Committee (REC reference 20-EMREC-007). Participants may find the tests, or receiving their scores, causes anxiety or stress. Previous exposure to similar tests may make this more familiar and reduce this anxiety. The study information will include signposting in case of distress. Study results will be disseminated to study participants, presented at conferences and published in a peer reviewed journal. No study participants will be identifiable in the study results.

Strengths and limitations

- Strength: Novel speech data collection in a preclinical cohort of participants at risk for dementia with AD biomarkers
- Strength: Study and materials designed with input from participants
- Strength: Collection of participant acceptability for test methodology
- Limitation: AD biomarkers not concurrently collected
- Limitation: Participants recruited from previous studies which are not diverse representations of wider population

For peer review only

Introduction

More than 40 million people worldwide are living with dementia [1]. There is a growing body of work to understand risk factors and how we can prevent dementia [2, 3]. Identification of biomarkers of Alzheimer's disease (AD) has been important in advancing the field [4]. Current gold standard biomarkers for AD include amyloid positron emission tomography (PET) scans [5], cerebrospinal fluid (CSF) amyloid and tau levels [6, 7] and blood based amyloid and tau concentrations [8]. However these techniques can be expensive [9], perceived as invasive or provoking anxiety for some participants [10], and may only be available in certain geographical locations.

Speech data is easy to collect and non-invasive, with recent studies identifying possible markers of AD [11]. Semantic errors and difficulties in semantic processing may be important in preclinical AD [12]. The failure to stop autocorrect errors when reading aloud correlates with A β 1-42 [13]. Changes in connected speech beginning as early as prodromal stages of the illness have been associated with autopsy confirmed AD, alongside linear declines in syntactic complexity and semantic content [14]. Acoustic features may also reveal information on emotional states and potential presence of neuropsychiatric symptoms such as apathy or depression which are comorbid conditions in AD [15-18].

Adapting traditional face to face cognitive assessments to allow remote data collection may increase access to both research studies and clinical services. Recent studies have shown that automated screening processes, including automated speech analysis, give as accurate neurocognitive screening decisions as human evaluators [19, 20]. These tasks are also validated in a telephone-based setting [21], Artificial Intelligence (AI) empowered pipelines can also capture clinically relevant features beyond that available in traditionally administered neurocognitive tests [22, 23].

The aim of the Speech on the Phone Assessment (SPeAk) is to collect speech data from participants across a spectrum of risk for AD, and assess the usefulness of speech features collected over the phone as potential digital biomarkers for AD.

Methods and analysis

Objectives

The objectives of the SPeAk study are:

- To develop algorithms to identify speech biomarkers in the data that are predictive of CSF AD biomarkers.
- To identify ceiling performance on a set of frequently performed cognitive tasks when administered on the telephone.
- To assess the test-retest reliability where the initial testing has a human element and the follow up testing is fully automated
- To explore participant acceptability of both with human and fully automated cognitive testing via the telephone.

- To conduct conversational analysis on the spontaneous speech generated in baseline testing session.

Participants: samples size and power calculation

Participants will all have engaged with in person cognitive testing at the research site within the last 12-24 months. The rationale for this is two-fold. Firstly, some of these participants were enrolled in the European Prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS) [24, 25] and have CSF biomarkers available for analysis in combination with the speech data collected in this protocol. Secondly, we wanted to recruit participants who were familiar with cognitive testing procedures to ensure they were psychologically robust (as cognitive testing can elicit anxiety) and had a reference point of comparison for acceptance evaluations.

The primary objective is the evaluation of an algorithm, and power calculations remain challenging in this field, with the general accepted wisdom being more data is better. Using the rationale explained by de la Fuente Garcia et al (2019) [26] a minimum of 75 participants would be sufficient for this analysis, however we will aim for at least 150 participants. Briefly in this protocol the authors justify the minimum sample size by placing the lower bounds between $1.2 \times f$ and $1.4 \times f$ where f is the number of features f the dataset. Using the Geneva Minimalistic Acoustic Parameter Set (GeMAPS) [27] which contains 62 features, the authors concluded a minimum of 75 participants would be required for 90% accuracy. In this study we anticipate 30-50 core features.

Experimental procedure and design

This study will trial Speech Analysis for cognitive assessments over the telephone in a study population previously exposed to in-person cognitive testing. Participants will be asked to complete two assessments during the study period. In the first testing session participants will interact with a staff member who will deliver the cognitive assessments on the phone, facilitated by an iPad application. These cognitive assessments will be audio recorded for speech analysis. There will also be free speech recorded during the telephone conversation which will be used for conversational analysis, and to assess the predictive power and clinical relevance of acoustic features extracted from spontaneous conversational speech [28]. In the follow up testing session, completed 3 months after the first session (+/- 14 days), participants will complete automated cognitive tests.

Consent

Participants will complete consent forms electronically using Online Survey software (<https://www.onlinesurveys.ac.uk/>). An electronic consent whereby participants will tick consent clauses and type their name has been deemed to be proportionate to the study risks. At the start of the first telephone call for the study, participants will be asked to verbally re-confirm consent to participate in the trial.

Demographics

Demographic details will be collected from all participants after they have provided informed consent to the study. Sex, age, education (years), living status and current medications will be recorded.

Spontaneous speech

1
2
3 Participants will be informed the rest of the session will be audio recorded and the session
4 will begin with approximately 5 minutes of spontaneous speech, prompted with the
5 completion of two conversational tasks. In the first task participants will be asked to respond
6 to the question: 'What animal makes the best pet?'. To encourage conversational exchange
7 between the interviewer and the participant, the interviewer will ask follow-up questions,
8 such as 'Why do you think that?' to elicit turn-taking. In task two participants will be invited
9 to engage in a game of 20 questions. The interviewer will explain briefly that it is a game in
10 which one participant thinks of an object, animal, plant or substance and the other participant
11 asks yes-or-no questions to determine what it is. To offer a level of standardisation the
12 interviewer will always think of an animal and the participant will ask questions to try and
13 guess the answer. The participant may only ask questions that can be answered with a yes or
14 a no. This data will only be collected at the baseline visit as a human interviewer is needed
15 for this.
16
17
18

19 *Cognitive testing*

20 Participants will complete a series of cognitive tasks selected to test a range of domains and
21 assessed for suitability to complete on the telephone. Encoding, immediate and delayed recall
22 of verbal words will be tested using the immediate and delayed Rey Auditory Visual
23 Learning Test (RAVLT) list learning task [29]. Between the initial and delayed recall tasks
24 participants will also be asked to complete the Wechsler Adult Intelligence Scale-IV (WAIS-
25 IV) forward digit span, phonemic (letter S) and semantic (category animals) verbal fluency
26 tasks. All test instructions will be read verbatim by a human interviewer at baseline and a
27 computerised voice at 3-month follow up. An audio recording is made of participant
28 responses which will be used for automatic transcription and scoring of the tasks.
29
30
31

32 *Results feedback*

33 Participants will be randomised 1:1 to receive feedback after each session or only after the
34 final session. This is to analyse whether participants have a preference for receiving
35 immediate feedback or are happy to receive feedback at the end of their involvement in the
36 study. Results will be sent to the participant using a proforma that explains the experimental
37 nature of the tasks, the expected score range for their age and their scores.
38
39

40 *Acceptability questionnaires*

41 Following each assessment point, participants will complete an acceptability questionnaire
42 designed for this study. These questionnaires will evaluate how easy the assessments were to
43 set up, how comfortable participants felt completing these assessments remotely, whether
44 participants preferred face to face assessments (completed in earlier trials), human tester
45 administered assessments or computer administered assessments and understand participants
46 experiences of receiving their test results.
47
48
49

50 *Previously acquired biomarker data*

51 Participants have previously taken part in studies which collected AD biomarkers (CSF, PET
52 and structural MRI), information about AD risk factors (such as apolipoprotein E genetic
53 status) and cognitive test batteries (Repeatable Battery for the Assessment of
54 Neuropsychological Status (RBANS), Clinical Dementia Rating scale (CDR), Mini Mental
55 State Examination (MMSE)). These data will be used in the machine learning model to build
56 the algorithm for the primary outcome.
57
58
59

60 **Data Management**

1
2
3 All speech data will be recorded on the Delta Testing App on an iPad and securely
4 transferred to a University of Edinburgh server. Participants are all assigned a unique ID code
5 which is used to label all of their study data. Due to the potential benefit of speech data for
6 future research, participants will be asked to consent to long term storage of their raw speech
7 files. These files will be securely held on University of Edinburgh servers and will be
8 available for analysis for the research team as part of this protocol, and in the future for
9 approved studies from researchers approved by the study team. Data from the EPAD LCS
10 study is held securely on the ADDF platform and relevant participant data will be extracted
11 for use in planned analyses. Demographic and questionnaire data will initially be stored using
12 onlinesurveys.ac.uk before transfer to the study server. All procedures are in line with the
13 University of Edinburgh data protection policy which is informed by the General Data
14 Protection Regulation (GDPR).
15
16

17 **Patient and public involvement statement**

18 Two participants from the EPAD LCS participant panel [30] were involved in the study
19 design process and reviewed the study documentation electronically. In particular the
20 participants were asked their opinions on whether they thought it was acceptable to conduct
21 cognitive assessments on the telephone and their thoughts on receiving results from the
22 assessments. Both participants agreed that completing cognitive assessments on the telephone
23 would be acceptable to them, and they felt that would apply to the rest of the group. Both
24 welcomed the feedback of results as this was something the panel had advocated for with
25 previous studies. The participants felt that it should be clear in the information sheet that
26 feedback would be available, so that if anyone who did not want to receive this would know
27 to decline the study invitation. All comments on the participant facing documentation were
28 incorporated into the final document to improve readability and maximise ease of use and
29 understanding of the study documents.
30
31
32
33

34 **Analysis**

35 To achieve the primary objective of the study, three different types of variables will be
36 computed and extracted from speech recordings: classical neuropsychological outcome
37 variables, novel or qualitative outcome variables based on produced language and low-level
38 speech descriptors. These measures will be computed on a task-level, meaning for each
39 cognitive task performed. A combination of predictive statistical and machine learning
40 models will be used, with patients' biomarker status as a target variable. Models will be
41 constructed with task specific features and using aggregated variables spanning multiple
42 cognitive tasks. Patients with missing data will not be considered in the final analysis. An
43 interim analysis will be performed once 100 participants have been recorded.
44
45
46

47 Secondary objectives are to evaluate ceiling effects and test-retest reliability. For this analysis
48 classical neuropsychological outcome variables will be extracted from the recorded audio. To
49 this end, human raters will use classical scoring schemes to score speech recordings of
50 cognitive tasks. This will lead to a dataset where each participant has multiple repeated
51 measurements of the same task. Results from multiple time points will be compared using a
52 repeated measures ANOVA and repeated measures correlation. Patients with missing data will
53 not be considered in the final analysis. An interim analysis will be performed once 100
54 participants have been recorded. We will use frequency statistics and qualitative analysis for
55 the acceptability questionnaires and compare baseline and 3-month responses used paired t-
56 tests. We will also analyse and predict levels of participant satisfaction, as recorded in the
57 acceptability questionnaires, based on acoustic and conversational speech features.
58
59
60

1
2
3 Analysis of the conversational (dialogue) data will generate speech interaction features and
4 vocalisation graphs which will be used as input representations for machine learning
5 modelling. These interactional features will be combined with low level acoustic descriptors
6 for predictive modelling of biomarker status and results of the cognitive tasks [31, 32].
7
8

9 **Ethics and dissemination**

10 This study has been reviewed and given favourable ethical opinion by the Edinburgh Medical
11 School Research Ethics Committee (EMREC) (REC Reference 20-EMREC-007). The study
12 raises few ethical concerns. Participants may find participation in cognitive testing causes
13 stress or anxiety. All participants invited to this study will previously have completed at least
14 one testing session in a face-to-face setting. We hope that this will reduce the risk of
15 participants experiencing stress as the cognitive tasks will be somewhat routine and
16 participants who did find the cognitive testing anxiety provoking in the preceding studies are
17 likely to decline this study invitation. Participants with concerns about their memory will be
18 encouraged to speak to their primary care practitioner or local Alzheimer's charity hotlines.
19
20

21 Receiving results from the cognitive tests may also lead participants to worry about their
22 memory. Providing participants with their results is increasingly common, and the plan to do
23 so in this study was well received by our participant panel advisors. Nearly half of
24 participants enrolled in a longitudinal AD study were interested to receive their cognitive test
25 results [33]. In that study participants highlighted receiving personal feedback would be a
26 motivation for engagement in longitudinal research studies [33]. We were not able to identify
27 studies reporting on participants' experiences of receiving their cognitive test results and will
28 evaluate this within the acceptability questionnaires in our study. Studies that have
29 investigated disclosure of AD risk factors, such as amyloid PET and APOE results, have
30 identified both benefits and harms of disclosure [34-36]. When done in a safe and appropriate
31 way, there is little short term psychological harm in risk factor disclosure [37]. On balance
32 disclosure, when done in a safe and sensible way, allows participants to plan and prepare
33 causing little psychological distress. As previously detailed, any participants who express
34 particular concern about their cognitive test scores will be encouraged to seek help from their
35 primary care practitioner or Alzheimer's disease charities.
36
37
38
39

40 The final ethical consideration is data security. Participants voices will be recorded, stored
41 and analysed. The study has been reviewed in line with university information governance
42 guidelines and complies with data security requirements. Publications arising from this study
43 will not identify any participants. Participants will be asked to explicitly consent to the
44 recording, storage and analysis of their voice data. Any participants who do not agree to this
45 will not be enrolled in the study.
46
47

48 **Conclusions**

49 The SPeAk study adds a unique opportunity to identify speech biomarkers in a preclinical
50 population at risk for dementia. The study will combine the collection of spontaneous speech
51 through two conversational tasks and a brief battery of cognitive tasks previously validated
52 in-person. The primary aim of the study is to develop an algorithm to identify speech features
53 which are predictive of the target variable of AD biomarkers (collected in a prior study).
54 Secondary objectives will evaluate test-retest reliability of the cognitive tasks when delivered
55 on the phone and participant acceptability of the assessments, including experiences of
56 receiving their cognitive test results. We will also undertake exploratory conversational
57 analysis of the spontaneous speech data.
58
59
60

1
2
3 Early detection of Alzheimer's disease is a priority for both clinical and research purposes.
4 Identifying a low-cost, non-invasive biomarker will be important in this field, and speech is
5 an exciting biomarker to explore.
6

7 **Acknowledgments**

8 We would like to thank the EPAD Scotland participant's panel for their review of the study,
9 and for all prospective participants for taking the time to consider this study.
10
11

12 **Financial disclosure/conflict of interest**

13 This work is supported by Janssen Pharmaceutica NV through a collaboration agreement
14 (Award/Grant number is not applicable). KL is an employee of Janssen.
15
16

17 **Author Statement**

18 All authors contributed to the design of the protocol. SG drafted the manuscript and all
19 authors reviewed and provided comments on drafts of the manuscript prior to submission.
20
21

22 **Data statement**

23 Data will be available through a data access process from the University of Edinburgh. This
24 data access will be managed by the study Principal Investigator (CWR), study coordinator
25 (SG) in conjunction with the University of Edinburgh data management specialists. As
26 speech data is potentially identifiable it is not appropriate for this data to be submitted to
27 open data access platforms external to the University of Edinburgh.
28
29

30 **References**

- 31 1. Nichols, E., et al., *Global, regional, and national burden of Alzheimer's disease and*
32 *other dementias, 1990-2016: a systematic analysis for the Global Burden of*
33 *Disease Study 2016*. The Lancet Neurology, 2019. **18**(1): p. 88-106.
- 34 2. Livingston, G., et al., *Dementia prevention, intervention, and care: 2020 report of the*
35 *Lancet Commission*. The Lancet, 2020. **396**(10248): p. 413-446.
- 36 3. Livingston, G., et al., *Dementia prevention, intervention, and care*. The Lancet, 2017.
37 **390**(10113): p. 2673-2734.
- 38 4. Humpel, C., *Identifying and validating biomarkers for Alzheimer's disease*. Trends in
39 biotechnology, 2011. **29**(1): p. 26-32.
- 40 5. Chételat, G., et al., *Amyloid-PET and ¹⁸F-FDG-PET in the diagnostic*
41 *investigation of Alzheimer's disease and other dementias*. The Lancet Neurology,
42 2020. **19**(11): p. 951-962.
- 43 6. Blennow, K., et al., *Predicting clinical decline and conversion to Alzheimer's disease*
44 *or dementia using novel Elecsys Aβ(1-42), pTau and tTau CSF immunoassays*.
45 Scientific Reports, 2019. **9**(1): p. 19024.
- 46 7. Blennow, K. and H. Zetterberg, *Biomarkers for Alzheimer's disease: current status*
47 *and prospects for the future*. Journal of Internal Medicine, 2018. **284**(6): p. 643-663.
- 48 8. Hampel, H., et al., *Blood-based biomarkers for Alzheimer disease: mapping the road*
49 *to the clinic*. Nature Reviews Neurology, 2018. **14**(11): p. 639-652.
- 50 9. Wittenberg, R., et al., *Economic impacts of introducing diagnostics for mild cognitive*
51 *impairment Alzheimer's disease patients*. Alzheimer's & dementia (New York, N. Y.),
52 2019. **5**: p. 382-387.
53
54
55
56
57
58
59
60

10. Peskind, E.R., et al., *Safety and acceptability of the research lumbar puncture*. Alzheimer Dis Assoc Disord, 2005. **19**(4): p. 220-5.
11. Boschi, V., et al., *Connected Speech in Neurodegenerative Language Disorders: A Review*. Front Psychol, 2017. **8**: p. 269.
12. Venneri, A., et al., *Diagnostic and prognostic role of semantic processing in preclinical Alzheimer's disease*. Biomarkers in Medicine, 2018. **12**(6): p. 637-651.
13. Gollan, T.H., et al., *Failure to stop autocorrect errors in reading aloud increases in aging especially with a positive biomarker for Alzheimer's disease*. Psychology and Aging, 2020. **35**(7): p. 1016-1025.
14. Ahmed, S., et al., *Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease*. Brain, 2013. **136**(12): p. 3727-3737.
15. Trevino, A.C., T.F. Quatieri, and N. Malyska, *Phonologically-based biomarkers for major depressive disorder*. EURASIP Journal on Advances in Signal Processing, 2011. **2011**(1): p. 42.
16. Yang, Y., C. Fairbairn, and J.F. Cohn, *Detecting Depression Severity from Vocal Prosody*. IEEE transactions on affective computing, 2013. **4**(2): p. 142-150.
17. Low, D.M., K.H. Bentley, and S.S. Ghosh, *Automated assessment of psychiatric disorders using speech: A systematic review*. Laryngoscope Investigative Otolaryngology, 2020. **5**(1): p. 96-116.
18. König, A., et al., *Detecting Apathy in Older Adults with Cognitive Disorders Using Automatic Speech Analysis*. J Alzheimers Dis, 2019. **69**(4): p. 1183-1193.
19. König, A., et al., *Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People*. Curr Alzheimer Res, 2018. **15**(2): p. 120-129.
20. König, A., et al., *Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task*. Dement Geriatr Cogn Disord, 2018. **45**(3-4): p. 198-209.
21. Tröger, J., et al., *Telephone-based Dementia Screening I: Automated Semantic Verbal Fluency Assessment*, in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2018, Association for Computing Machinery: New York, NY, USA. p. 59–66.
22. Linz, N., et al., *Using Neural Word Embeddings in the Analysis of the Clinical Semantic Verbal Fluency Task*. IWCS 2017 — 12th International Conference on Computational Semantics — Short papers, 2017.
23. Linz, N., et al. *Predicting Dementia Screening and Staging Scores from Semantic Verbal Fluency Performance*. in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. 2017.
24. Ritchie, C.W., et al., *The European Prevention of Alzheimer's Dementia (EPAD) Longitudinal Cohort Study: Baseline Data Release V500.0*. The Journal of Prevention of Alzheimer's Disease, 2020. **7**(1): p. 8-13.
25. Ritchie, C.W., et al., *Development of interventions for the secondary prevention of Alzheimer's dementia: the European Prevention of Alzheimer's Dementia (EPAD) project*. The Lancet Psychiatry, 2016. **3**(2): p. 179-186.
26. de la Fuente Garcia, S., C.W. Ritchie, and S. Luz, *Protocol for a conversation-based analysis study: PREVENT-ED investigates dialogue features that may help predict dementia onset in later life*. BMJ Open, 2019. **9**(3): p. e026254.

- 1
- 2
- 3
- 4 27. Eyben, F., et al., *The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice*
- 5 *Research and Affective Computing*. IEEE Transactions on Affective Computing, 2016.
- 6 **7(2)**: p. 190-202.
- 7 28. Haider, F., S.d.l. Fuente, and S. Luz, *An Assessment of Paralinguistic Acoustic Features*
- 8 *for Detection of Alzheimer's Dementia in Spontaneous Speech*. IEEE Journal of
- 9 Selected Topics in Signal Processing, 2020. **14(2)**: p. 272-281.
- 10 29. Rey, A., *L'examen clinique en psychologie. [The clinical examination in psychology.]*.
- 11 *L'examen clinique en psychologie*. 1958, Oxford, England: Presses Universitaires De
- 12 France. 222-222.
- 13 30. Gregory, S., et al., *Involving research participants in a pan-European research*
- 14 *initiative: the EPAD participant panel experience*. Research Involvement and
- 15 Engagement, 2020. **6(1)**: p. 62.
- 16 31. Luz, S. and S. De La Fuente, *A method for analysis of patient speech in dialogue for*
- 17 *dementia detection*. , in *Proceedings of LREC 2018*, D. Kokkinakis, Editor. 2018,
- 18 European Language Resources Association: Paris. p. 7-12.
- 19 32. Haider, F., S. De La Fuente, and S. Luz, *An assessment of paralinguistic acoustic*
- 20 *features for detection of Alzheimer's Dementia in spontaneous speech*. IEEE Journal
- 21 of Selected Topics in Signal Processing, 2019. **14(2)**: p. 272-281.
- 22 33. Gooblar, J., et al., *Attitudes of Research Participants and the General Public*
- 23 *Regarding Disclosure of Alzheimer Disease Research Results*. JAMA neurology, 2015.
- 24 **72(12)**: p. 1484-1490.
- 25 34. Vanderschaeghe, G., et al., *Amnesic MCI patients' experiences after disclosure of*
- 26 *their amyloid PET result in a research context*. Alzheimer's Research & Therapy,
- 27 2017. **9(1)**: p. 92.
- 28 35. Largent, E.A., et al., *Cognitively unimpaired adults' reactions to disclosure of amyloid*
- 29 *PET scan results*. PLOS ONE, 2020. **15(2)**: p. e0229137.
- 30 36. Christensen, K.D., et al., *Disclosing genetic risk for Alzheimer's dementia to*
- 31 *individuals with mild cognitive impairment*. Alzheimer's & Dementia: Translational
- 32 Research & Clinical Interventions, 2020. **6(1)**: p. e12002.
- 33 37. Grill, J.D., et al., *Short-term Psychological Outcomes of Disclosing Amyloid Imaging*
- 34 *Results to Research Participants Who Do Not Have Cognitive Impairment*. JAMA
- 35 Neurology, 2020. **77(12)**: p. 1504-1513.
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

BMJ Open

Remote data collection speech analysis and prediction of the identification of Alzheimer's disease biomarkers in people at risk for Alzheimer's disease dementia and: the Speech on the Phone Assessment (SPeAk) prospective observational study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-052250.R1
Article Type:	Protocol
Date Submitted by the Author:	26-Nov-2021
Complete List of Authors:	Gregory, Sarah; The University of Edinburgh Centre for Clinical Brain Sciences Linz, Nicklas; German Research Centre for Artificial Intelligence Saarbrücken Branch König, Alexandra; National Institute for Research in Computer Science and Automation Langel, Kai; Janssen Healthcare Innovation Pullen, Hannah; The University of Edinburgh Centre for Clinical Brain Sciences Luz, Saturnino; The University of Edinburgh Usher Institute of Population Health Sciences and Informatics Harrison, John; Metis Cognition Ltd; Vrije Universiteit Amsterdam, Department of Neurology Ritchie, Craig; The University of Edinburgh Centre for Clinical Brain Sciences
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Neurology
Keywords:	Dementia < NEUROLOGY, Delirium & cognitive disorders < PSYCHIATRY, Old age psychiatry < PSYCHIATRY

SCHOLARONE™
Manuscripts

1
2
3 Remote data collection speech analysis and prediction of the identification of
4 Alzheimer's disease biomarkers in people at risk for Alzheimer's disease
5 dementia and: the Speech on the Phone Assessment (SPeAk) prospective
6 observational study
7
8

9 Sarah Gregory¹, Nicklas Linz², Alexandra König³, Kai Langel⁴, Hannah Pullen¹, Saturnino
10 Luz⁵, John Harrison^{6, 7, 8}, Craig Ritchie¹
11
12

13 ¹ Edinburgh Dementia Prevention, Centre for Clinical Brain Sciences, University of
14 Edinburgh, Edinburgh, United Kingdom

15 ² KI Elements, Saarbrücken, Germany

16 ³ National Institute for Research in Computer Science and Automation (INRIA), France

17 ⁴ Janssen Clinical Innovation, Spain

18 ⁵ Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

19 ⁶ Metis Cognition Ltd, Kilmington Common, United Kingdom

20 ⁷ Alzheimer Center Amsterdam, Department of Neurology, Amsterdam UMC, Vrije
21 Universiteit, Amsterdam, The Netherlands

22 ⁸ Institute of Psychiatry, Psychology & Neuroscience, King's College London, United
23 Kingdom
24
25

26
27 Email addresses: Sarah.Gregory@ed.ac.uk; nicklas@ki-elements.de; alexandra@ki-
28 elements.de; KLangel1@its.jnj.com; Hannah.Pullen@ed.ac.uk; s.luz@ed.ac.uk;
29 johnnpc@btinternet.com; craig.ritchie@ed.ac.uk
30
31

32 Corresponding author: Sarah Gregory, Edinburgh Dementia Prevention, Biocube 1, 9 Little
33 France Road, Edinburgh, EH16 4UX; Sarah.Gregory@ed.ac.uk
34

35 Keywords: speech, biomarkers, Alzheimer's disease, cognition
36

37 Word count: 3595 words
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Introduction: Identifying cost effective, non-invasive biomarkers of Alzheimer's disease (AD) is a clinical and research priority. Speech data is easy to collect, and studies suggest it can identify those with AD. We do not know if speech features can predict AD biomarkers in a preclinical population.

Methods and analysis: The Speech on the Phone Assessment (SPeAk) study is a prospective observational study. SPeAk recruits participants aged 50 years and over who have previously completed studies with AD biomarker collection. Participants complete a baseline telephone assessment including spontaneous speech and cognitive tests. A 3-month visit will repeat the cognitive tests with a conversational artificial intelligence bot. Participants complete acceptability questionnaires after each visit. Participants are randomised to receive their cognitive test results either after each visit or only after they have completed the study. We will combine SPeAK data with AD biomarker data collected in a previous study and analyse for correlations between extracted speech features and AD biomarkers. The outcome of this analysis will inform the develop of an algorithm for prediction of AD risk based on speech features.

Ethics and dissemination: This study has been approved by the Edinburgh Medical School Research Ethics Committee (REC reference 20-EMREC-007). Participants may find the tests, or receiving their scores, causes anxiety or stress. Previous exposure to similar tests may make this more familiar and reduce this anxiety. The study information will include signposting in case of distress. Study results will be disseminated to study participants, presented at conferences and published in a peer reviewed journal. No study participants will be identifiable in the study results.

Strengths and limitations

- A key strength of this study is the novel speech data collection in a preclinical cohort of participants at risk for dementia with AD biomarkers.
- The SPeAk study and associated materials have been designed with input from participants to ensure the project is of interest and value to the target participant population.
- This study will collect data on participant acceptability of the test methodology to collect critical information for future projects.
- Unfortunately in this project there is no concurrent collection of AD biomarkers so only historical data is available.
- The participants recruited to this study do not represent the diversity of the wider population as the previous studies they were involved did not have inclusive recruitment strategies.

Introduction

More than 40 million people worldwide are living with dementia [1]. There is a growing body of work to understand risk factors and how we can prevent dementia [2, 3]. Identification of biomarkers of Alzheimer's disease (AD) has been important in advancing the field [4]. Current gold standard biomarkers for AD include amyloid positron emission tomography (PET) scans [5], cerebrospinal fluid (CSF) amyloid and tau levels [6, 7] and blood based amyloid and tau concentrations [8]. However these techniques can be expensive [9], perceived as invasive or provoking anxiety for some participants [10], and may only be available in certain geographical locations.

Speech data is easy to collect and non-invasive, with recent studies identifying possible markers of AD [11]. Semantic errors and difficulties in semantic processing may be important in preclinical AD [12]. The failure to stop autocorrect errors when reading aloud correlates with A β 1-42 [13]. Changes in connected speech beginning as early as prodromal stages of the illness have been associated with autopsy confirmed AD, alongside linear declines in syntactic complexity and semantic content [14]. Participants with preclinical AD, defined by amyloid beta positivity on a PET scan, were found to have more rapid declines on specific word content, a feature of connected speech [15]. Acoustic features may also reveal information on emotional states and potential presence of neuropsychiatric symptoms such as apathy or depression which are comorbid conditions in AD [16-19].

Adapting traditional face to face cognitive assessments to allow remote data collection may increase access to both research studies and clinical services. Recent studies have shown that automated screening processes, including automated speech analysis, give as accurate neurocognitive screening decisions as human evaluators [20-22]. These tasks are also validated in a telephone-based setting [23], Artificial Intelligence (AI) empowered pipelines can also capture clinically relevant features beyond that available in traditionally administered neurocognitive tests [24, 25].

The aim of the Speech on the Phone Assessment (SPeAk) is to collect speech data from participants across a spectrum of risk for AD, and assess the usefulness of speech features collected over the phone as potential digital biomarkers for AD.

Methods and analysis

Study design

The SPeAk study is a prospective observational study.

Objectives

The objectives of the SPeAk study are:

- To develop algorithms to identify speech biomarkers in the data that are predictive of CSF AD biomarkers.
- To identify ceiling performance on a set of frequently performed cognitive tasks when administered on the telephone.

- To assess the test-retest reliability where the initial testing has a human element and the follow up testing is fully automated
- To explore participant acceptability of both with human and fully automated cognitive testing via the telephone.
- To conduct conversational analysis on the spontaneous speech generated in baseline testing session.

The SPeAK project began in January 2021 and is expected to complete data collection December 2021, with data analysis ongoing until June 2022.

Participants: samples size and power calculation

Participants will all have engaged with in person cognitive testing at the research site within the last 12-24 months. The rationale for this is two-fold. Firstly, some of these participants were enrolled in the European Prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS) [26, 27] and have CSF biomarkers available for analysis in combination with the speech data collected in this protocol. Participants formerly enrolled in the EPAD LCS represent a range of risk states for AD dementia, from cognitively healthy to preclinical AD. Participants recruited from other studies are also either cognitively healthy or have preclinical AD. Importantly none of the participants are currently living with dementia. Secondly, we wanted to recruit participants who were familiar with cognitive testing procedures to ensure they were psychologically robust (as cognitive testing can elicit anxiety) and had a reference point of comparison for acceptance evaluations. All participants included are known to be native or fluent English language speakers from their previous study involvement. If fluent but non-native speakers chose to enroll in the SPeAk study a sensitivity analysis will be conducted to assess for any impact of non-native language patterns.

Participants will be included in this study if they have previously enrolled in the EPAD LCS or another closed commercial research study at the Edinburgh Dementia Prevention research site. Participants must have a landline or mobile phone and have the capacity to consent to the study. Participants will be excluded from the study if they do not have the capacity to provide informed consent or if they anticipate that they will not be available by landline or mobile phone at the three month follow up time point. As participants are being recruited from previous studies we have not specified further eligibility criteria. On the basis of the previous studies all participants will be aged 50 years or above, cognitively healthy or with mild cognitive impairment, will not have dementia, will be fluent English speakers and will not have visual or hearing impairments severe enough to interfere with cognitive testing.

The primary objective is the evaluation of an algorithm, and power calculations remain challenging in this field, with the general accepted wisdom being more data is better. Using the rationale explained by de la Fuente Garcia et al (2019) [28] a minimum of 75 participants would be sufficient for this analysis, however we will aim for at least 150 participants. Briefly in this protocol the authors justify the minimum sample size by placing the lower bounds between $1.2 \times f$ and $1.4 \times f$ where f is the number of features of the dataset. Using the Geneva Minimalistic Acoustic Parameter Set (GeMAPS) [29] which contains 62 features, the authors concluded a minimum of 75 participants would be required for 90% accuracy. In this study we anticipate 30-50 core features. For the machine learning analysis we have adopted a pragmatic approach to sample size calculation, with an aim of 150 participants allowing us to develop two reasonably sized datasets, one for learning and one for testing. A simple classification algorithm such as the Euclidean distance requires $1.2 \times v$ where v is the number

of variables. If we anticipate 50 core speech features as well as a small number of demographic covariates (described later) we arrive at a required minimum sample size of 66.

Experimental procedure and design

This study will trial Speech Analysis for cognitive assessments over the telephone in a study population previously exposed to in-person cognitive testing. Participants will be asked to complete two assessments during the study period. In the first testing session participants will interact with a staff member who will deliver the cognitive assessments on the phone, facilitated by an iPad application. The participants do not need to have an iPad themselves, the app will call either a landline or mobile phone number. These cognitive assessments will be audio recorded for speech analysis. There will also be free speech recorded during the telephone conversation which will be used for conversational analysis, and to assess the predictive power and clinical relevance of acoustic features extracted from spontaneous conversational speech [30]. In the follow up testing session, completed 3 months after the first session (+/- 14 days), participants will complete automated cognitive tests. For the automated tests a computerised voice will read the instructions and initiate tasks for the participants. A computer voice was selected instead of a prerecorded human voice to ensure it was clear to participants that they are engaging with a computer and can not freely interact as with a human tester.

Consent

Participants will complete consent forms electronically using Online Survey software (<https://www.onlinesurveys.ac.uk/>). An electronic consent whereby participants will tick consent clauses and type their name has been deemed to be proportionate to the study risks. At the start of the first telephone call for the study, participants will be asked to verbally re-confirm consent to participate in the trial. If the participant does not have the capacity to consent to their participation in the study they will not continue in the study. Members of the research team are experienced in working with this population and assessing capacity. Only appropriately trained staff will be delegated to this task. Senior staff will be available to discuss any concerns about a participant's capacity to consent.

Demographics

Demographic details will be collected from all participants after they have provided informed consent to the study. Sex, age, education (years), living status and current medications will be recorded.

Preparing the participant

Prior to the spontaneous speech and cognitive tests the participants will be prepared for the session by the interviewer. The interviewer will check the audio quality, both for the participant and for the interviewer. The participants will be asked to position themselves in a quiet room away from distractions and to ensure they do not have anything available to write on (such as pen and paper or an electronic device). Participants will be informed that if the call is terminated for any reason an attempt will be made to re-establish the call. Prior to the second call, which is led by the computer, the participants will be reminded of these preparations in an email. If participants would like a phone call to review these preparations this will be offered.

Spontaneous speech

1
2
3 Participants will be informed the rest of the session will be audio recorded and the session
4 will begin with approximately 5 minutes of spontaneous speech, prompted with the
5 completion of two conversational tasks. In the first task participants will be asked to respond
6 to the question: 'What animal makes the best pet?'. To encourage conversational exchange
7 between the interviewer and the participant, the interviewer will ask follow-up questions,
8 such as 'Why do you think that?' to elicit turn-taking. In task two participants will be invited
9 to engage in a game of 20 questions. The interviewer will explain briefly that it is a game in
10 which one participant thinks of an object, animal, plant or substance and the other participant
11 asks yes-or-no questions to determine what it is. To offer a level of standardisation the
12 interviewer will always think of the same animal and the participant will ask questions to try
13 and guess the answer. The participant may only ask questions that can be answered with a yes
14 or a no. This data will only be collected at the baseline visit as a human interviewer is needed
15 for this.
16
17
18

19 *Cognitive testing*

20 Participants will complete a series of cognitive tasks selected to test a range of domains and
21 assessed for suitability to complete on the telephone. Domains were selected to represent
22 cognitive abilities sensitive to decline in early AD [31, 32]. Encoding, immediate and delayed
23 recall of verbal words will be tested using the immediate and delayed Rey Auditory Visual
24 Learning Test (RAVLT) list learning task [33]. Between the initial and delayed recall tasks
25 participants will also be asked to complete the Wechsler Adult Intelligence Scale-IV (WAIS-
26 IV) forward digit span, phonemic (letter S) and semantic (category animals) verbal fluency
27 tasks. All test instructions will be read verbatim by a human interviewer at baseline and a
28 computerised voice at 3-month follow up. An audio recording is made of participant
29 responses which will be used for automatic transcription and scoring of the tasks.
30
31
32

33 *Results feedback*

34 Participants will be randomised 1:1 to receive results of their cognitive tests after each
35 session or only after the final session. This is to analyse whether participants have a
36 preference for receiving immediate feedback or are happy to receive feedback at the end of
37 their involvement in the study. No individual level speech feature data will be feedback to
38 participants as this remains experimental. Participants will be provided with a copy of the
39 overall study results. Individual cognitive test results will be sent to the participant using a
40 proforma that explains the experimental nature of the tasks (i.e. being completed on the
41 phone rather than face to face), the expected score range for their age (based on published
42 normative range for face to face assessment) and their scores. This proforma will be sent via
43 email from the research team. If a participant has a score that falls below the published
44 normative range their data will be reviewed by the Chief Investigator (CWR, an experienced
45 Consultant Psychiatrist, clinical triallist and Professor of Old Age Psychiatry) prior to result
46 feedback. Any results that may be of concern will be highlighted to the participant in a phone
47 call prior to receiving the email with their results and a letter to their GP written if the
48 participant consents to this.
49
50
51

52 *Acceptability questionnaires*

53 Following each assessment point, participants will complete an acceptability questionnaire
54 designed for this study. These questionnaires will evaluate how easy the assessments were to
55 set up, how comfortable participants felt completing these assessments remotely, whether
56 participants preferred face to face assessments (completed in earlier trials), human tester
57 administered assessments or computer administered assessments and understand participants
58 experiences of receiving their test results.
59
60

Previously acquired biomarker data

Participants have previously taken part in studies which collected AD biomarkers (CSF, PET and structural MRI), information about AD risk factors (such as apolipoprotein E genetic status) and cognitive test batteries (Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), Clinical Dementia Rating scale (CDR), Mini Mental State Examination (MMSE)). These data will be used in the machine learning model to build the algorithm for the primary outcome.

Data Management

All speech data will be recorded on the Delta Testing App on an iPad and securely transferred to a University of Edinburgh server. Participants are all assigned a unique ID code which is used to label all of their study data. Due to the potential benefit of speech data for future research, participants will be asked to consent to long term storage of their raw speech files. These files will be securely held on University of Edinburgh servers and will be available for analysis for the research team as part of this protocol, and in the future for approved studies from researchers approved by the study team. Data from the EPAD LCS study are held securely on the Alzheimer's Drug Discovery Foundation (ADDF) platform and relevant participant data will be extracted for use in planned analyses. Demographic and questionnaire data will initially be stored using onlinesurveys.ac.uk before transfer to the study server. All procedures are in line with the University of Edinburgh data protection policy which is informed by the General Data Protection Regulation (GDPR).

Patient and public involvement statement

Two participants from the EPAD LCS participant panel [34] were involved in the study design process and reviewed the study documentation electronically. In particular the participants were asked their opinions on whether they thought it was acceptable to conduct cognitive assessments on the telephone and their thoughts on receiving results from the assessments. Both participants agreed that completing cognitive assessments on the telephone would be acceptable to them, and they felt that would apply to the rest of the group. Both welcomed the feedback of results as this was something the panel had advocated for with previous studies. The participants felt that it should be clear in the information sheet that feedback would be available, so that if anyone who did not want to receive this would know to decline the study invitation. All comments on the participant facing documentation were incorporated into the final document to improve readability and maximise ease of use and understanding of the study documents.

Analysis

To achieve the primary objective of the study, three different types of variables will be computed and extracted from speech recordings: classical neuropsychological outcome variables, novel or qualitative outcome variables based on produced language and low-level speech descriptors. These measures will be computed on a task-level, meaning for each cognitive task performed. Language and speech variables will be extracted on a linguistic and acoustic level based. These include temporal, voice source, formant, semantic, and syntactic variables. Some variables are task specific and encode performance or strategy related to a task, some are general descriptors of voice quality. For the primary analysis we will develop machine learning models, with patients' biomarker status as a target variable. Models will be constructed with task specific features and using aggregated variables spanning multiple

1
2
3 cognitive tasks. Patients with missing data will not be considered in the final analysis. An
4 interim analysis will be performed once 100 participants have been recorded.
5 Secondary objectives are to evaluate ceiling effects and stability in cognitive performance
6 between a human and a computer tester. For this analysis classical neuropsychological outcome
7 variables will be extracted from the recorded audio. To this end, human raters will use classical
8 scoring schemes to score speech recordings of cognitive tasks. This will lead to a dataset where
9 each participant has multiple repeated measurements of the same task. Results from multiple
10 time points will be compared using a repeated measures ANOVA and repeated measures
11 correlation. Patients with missing data will not be considered in the final analysis. An interim
12 analysis will be performed once 100 participants have been recorded. We will use frequency
13 statistics and qualitative analysis for the acceptability questionnaires and compare baseline and
14 3-month responses used paired t-tests. We will also analyse and predict levels of participant
15 satisfaction, as recorded in the acceptability questionnaires, based on acoustic and
16 conversational speech features.
17
18
19

20
21 Analysis of the conversational (dialogue) data will generate speech interaction features and
22 vocalisation graphs which will be used as input representations for machine learning
23 modelling. These interactional features will be combined with low level acoustic descriptors
24 for predictive modelling of biomarker status and results of the cognitive tasks [35, 36].
25

26 **Ethics and dissemination**

27 This study has been reviewed and given favourable ethical opinion by the Edinburgh Medical
28 School Research Ethics Committee (EMREC) (REC Reference 20-EMREC-007). The study
29 raises few ethical concerns. Participants may find participation in cognitive testing causes
30 stress or anxiety. All participants invited to this study will previously have completed at least
31 one testing session in a face-to-face setting. We hope that this will reduce the risk of
32 participants experiencing stress as the cognitive tasks will be somewhat routine and
33 participants who did find the cognitive testing anxiety provoking in the preceding studies are
34 likely to decline this study invitation. Participants with concerns about their memory will be
35 encouraged to speak to their primary care practitioner or local Alzheimer's charity hotlines.
36
37
38

39 All participants will provide informed consent prior to participating in the SPeAk study. We
40 will use electronic consent for this study, via the Online Survey software
41 (<https://www.onlinesurveys.ac.uk/>). This software provides an electronic consent whereby
42 participants will tick consent clauses and type their name. Electronic consent has been
43 deemed to be proportionate to the SPeAk study risks. At the start of the first telephone call
44 for the study, participants will be asked to verbally re-confirm consent to participate in the
45 trial. If the participant does not have the capacity to consent to their participation in the
46 study they will not continue in the study. Members of the research team are experienced in working
47 with this population and assessing capacity. Only appropriately trained staff will be delegated
48 to this task. Senior staff will be available to discuss any concerns about a participant's
49 capacity to consent.
50
51

52
53 Receiving results from the cognitive tests may also lead participants to worry about their
54 memory. Providing participants with their results is increasingly common, and the plan to do
55 so in this study was well received by our participant panel advisors. Nearly half of
56 participants enrolled in a longitudinal AD study were interested to receive their cognitive test
57 results [37]. In that study participants highlighted receiving personal feedback would be a
58 motivation for engagement in longitudinal research studies [37]. We were not able to identify
59 studies reporting on participants' experiences of receiving their cognitive test results and will
60

1
2
3 evaluate this within the acceptability questionnaires in our study. Studies that have
4 investigated disclosure of AD risk factors, such as amyloid PET and APOE results, have
5 identified both benefits and harms of disclosure [38-40]. When done in a safe and appropriate
6 way, there is little short term psychological harm in risk factor disclosure [41]. On balance
7 disclosure, when done in a safe and sensible way, allows participants to plan and prepare
8 causing little psychological distress. As previously detailed, any participants who express
9 particular concern about their cognitive test scores will be encouraged to seek help from their
10 primary care practitioner or Alzheimer's disease charities.
11
12

13 The final ethical consideration is data security. Participants voices will be recorded, stored
14 and analysed. The study has been reviewed in line with university information governance
15 guidelines and complies with data security requirements. Publications arising from this study
16 will not identify any participants. Participants will be asked to explicitly consent to the
17 recording, storage and analysis of their voice data. Any participants who do not agree to this
18 will not be enrolled in the study.
19
20

21 **Discussion**

22 The SPeAk study adds a unique opportunity to identify speech biomarkers in a preclinical
23 population at risk for dementia. The study will combine the collection of spontaneous speech
24 through two conversational tasks and a brief battery of cognitive tasks previously validated
25 in-person. The primary aim of the study is to develop an algorithm to identify speech features
26 which are predictive of the target variable of AD biomarkers (collected in a prior study).
27 Secondary objectives will evaluate test-retest reliability of the cognitive tasks when delivered
28 on the phone and participant acceptability of the assessments, including experiences of
29 receiving their cognitive test results. We will also undertake exploratory conversational
30 analysis of the spontaneous speech data.
31
32
33

34 Early detection of Alzheimer's disease is a priority for both clinical and research purposes.
35 Identifying a low-cost, non-invasive biomarker will be important in this field, and speech is
36 an exciting biomarker to explore.
37
38

39 **Acknowledgments**

40 We would like to thank the EPAD Scotland participant's panel for their review of the study,
41 and for all prospective participants for taking the time to consider this study.
42
43

44 **Funding**

45 This work is supported by Janssen Pharmaceutica NV through a collaboration agreement
46 (Award/Grant number is not applicable).
47
48

49 **Competing interests**

50 KL is an employee of Janssen (the funder). NL & AK are employees of KI Elements who
51 own and provide the Delta Testing App. Dr. Harrison reports receipt of personal fees in the
52 past 2 years from Actinogen, AlzeCure, Aptinyx, Astra Zeneca, Athira Therapeutics, Axon
53 Neuroscience, Axovant, Bial Biotech, Biogen Idec, BlackThornRx, Boehringer Ingelheim,
54 Brands2life, Cerecin, Cognito, Cognition Therapeutics, Compass Pathways, Corlieve,
55 Curasen, EIP Pharma, Eisai, G4X Discovery, GfHEU, Heptares, Ki Elements, Lundbeck,
56 Lysosome Therapeutics, MyCognition, Neurocentria, Neurocog, Neurodyn Inc, Neurotrack,
57 the NHS, Novartis, Novo Nordisk, Nutricia, Probiobdrug, Prothema, Recognify, Regeneron,
58 reMYND, Rodin Therapeutics, Samumed, Sanofi, Signant, Syndesi Therapeutics, Takeda,
59 Vivoryon Therapeutics and Winterlight Labs. Additionally, he holds stock options in
60

1
2
3 Neurotrack Inc. and is a joint holder of patents with My Cognition Ltd. CWR has received
4 consultancy fees from Biogen, Eisai, MSD, Actinogen, Roche, and Eli Lilly, as well as
5 payment or honoraria from Roche and Eisai. SG, HP & SL have no conflicts to declare.
6
7

8 **Author Statement**

9 SG coordinated the design of the protocol, managed the ethical approval process, drafted the
10 manuscript and edited the final version. NL, AK, KL, HP, SL and JL all contributed to the
11 design of the protocol and provided comments on the manuscript drafts. CWR is the principal
12 investigator of the protocol and oversaw the design and ethical approval process as well as
13 provided comments on the manuscript drafts.
14
15
16

17 **Data statement**

18 Data will be available through a data access process from the University of Edinburgh. This
19 data access will be managed by the study Principal Investigator (CWR), study coordinator
20 (SG) in conjunction with the University of Edinburgh data management specialists. As
21 speech data is potentially identifiable it is not appropriate for this data to be submitted to
22 open data access platforms external to the University of Edinburgh.
23
24

25 **References**

- 26
- 27
- 28 1. Nichols, E., et al., *Global, regional, and national burden of Alzheimer's disease and*
29 *other dementias, 1990–2016: a systematic analysis for the Global Burden of*
30 *Disease Study 2016*. The Lancet Neurology, 2019. **18**(1): p. 88-106.
- 31 2. Livingston, G., et al., *Dementia prevention, intervention, and care: 2020 report of the*
32 *Lancet Commission*. The Lancet, 2020. **396**(10248): p. 413-446.
- 33 3. Livingston, G., et al., *Dementia prevention, intervention, and care*. The Lancet, 2017.
34 **390**(10113): p. 2673-2734.
- 35 4. Humpel, C., *Identifying and validating biomarkers for Alzheimer's disease*. Trends in
36 biotechnology, 2011. **29**(1): p. 26-32.
- 37 5. Chételat, G., et al., *Amyloid-PET and ¹⁸F-FDG-PET in the diagnostic*
38 *investigation of Alzheimer's disease and other dementias*. The Lancet Neurology,
39 2020. **19**(11): p. 951-962.
- 40 6. Blennow, K., et al., *Predicting clinical decline and conversion to Alzheimer's disease*
41 *or dementia using novel Elecsys Aβ(1–42), pTau and tTau CSF immunoassays*.
42 Scientific Reports, 2019. **9**(1): p. 19024.
- 43 7. Blennow, K. and H. Zetterberg, *Biomarkers for Alzheimer's disease: current status*
44 *and prospects for the future*. Journal of Internal Medicine, 2018. **284**(6): p. 643-663.
- 45 8. Hampel, H., et al., *Blood-based biomarkers for Alzheimer disease: mapping the road*
46 *to the clinic*. Nature Reviews Neurology, 2018. **14**(11): p. 639-652.
- 47 9. Wittenberg, R., et al., *Economic impacts of introducing diagnostics for mild cognitive*
48 *impairment Alzheimer's disease patients*. Alzheimer's & dementia (New York, N. Y.),
49 2019. **5**: p. 382-387.
- 50 10. Peskind, E.R., et al., *Safety and acceptability of the research lumbar puncture*.
51 Alzheimer Dis Assoc Disord, 2005. **19**(4): p. 220-5.
- 52 11. Boschi, V., et al., *Connected Speech in Neurodegenerative Language Disorders: A*
53 *Review*. Front Psychol, 2017. **8**: p. 269.
54
55
56
57
58
59
60

12. Venneri, A., et al., *Diagnostic and prognostic role of semantic processing in preclinical Alzheimer's disease*. *Biomarkers in Medicine*, 2018. **12**(6): p. 637-651.
13. Gollan, T.H., et al., *Failure to stop autocorrect errors in reading aloud increases in aging especially with a positive biomarker for Alzheimer's disease*. *Psychology and Aging*, 2020. **35**(7): p. 1016-1025.
14. Ahmed, S., et al., *Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease*. *Brain*, 2013. **136**(12): p. 3727-3737.
15. Mueller, K.D., et al., *Amyloid beta associations with connected speech in cognitively unimpaired adults*. *Alzheimers Dement (Amst)*, 2021. **13**(1): p. e12203.
16. Trevino, A.C., T.F. Quatieri, and N. Malyska, *Phonologically-based biomarkers for major depressive disorder*. *EURASIP Journal on Advances in Signal Processing*, 2011. **2011**(1): p. 42.
17. Yang, Y., C. Fairbairn, and J.F. Cohn, *Detecting Depression Severity from Vocal Prosody*. *IEEE transactions on affective computing*, 2013. **4**(2): p. 142-150.
18. Low, D.M., K.H. Bentley, and S.S. Ghosh, *Automated assessment of psychiatric disorders using speech: A systematic review*. *Laryngoscope Investigative Otolaryngology*, 2020. **5**(1): p. 96-116.
19. König, A., et al., *Detecting Apathy in Older Adults with Cognitive Disorders Using Automatic Speech Analysis*. *J Alzheimers Dis*, 2019. **69**(4): p. 1183-1193.
20. König, A., et al., *Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People*. *Curr Alzheimer Res*, 2018. **15**(2): p. 120-129.
21. König, A., et al., *Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task*. *Dement Geriatr Cogn Disord*, 2018. **45**(3-4): p. 198-209.
22. Themistocleous, C., M. Eckerström, and D. Kokkinakis, *Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls*. *PLOS ONE*, 2020. **15**(7): p. e0236009.
23. Tröger, J., et al., *Telephone-based Dementia Screening I: Automated Semantic Verbal Fluency Assessment*, in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2018, Association for Computing Machinery: New York, NY, USA. p. 59–66.
24. Linz, N., et al., *Using Neural Word Embeddings in the Analysis of the Clinical Semantic Verbal Fluency Task*. *IWCS 2017 — 12th International Conference on Computational Semantics — Short papers*, 2017.
25. Linz, N., et al. *Predicting Dementia Screening and Staging Scores from Semantic Verbal Fluency Performance*. in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. 2017.
26. Ritchie, C.W., et al., *The European Prevention of Alzheimer's Dementia (EPAD) Longitudinal Cohort Study: Baseline Data Release V500.0*. *The Journal of Prevention of Alzheimer's Disease*, 2020. **7**(1): p. 8-13.
27. Ritchie, C.W., et al., *Development of interventions for the secondary prevention of Alzheimer's dementia: the European Prevention of Alzheimer's Dementia (EPAD) project*. *The Lancet Psychiatry*, 2016. **3**(2): p. 179-186.
28. de la Fuente Garcia, S., C.W. Ritchie, and S. Luz, *Protocol for a conversation-based analysis study: PREVENT-ED investigates dialogue features that may help predict dementia onset in later life*. *BMJ Open*, 2019. **9**(3): p. e026254.

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
29. Eyben, F., et al., *The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing*. IEEE Transactions on Affective Computing, 2016. **7**(2): p. 190-202.
30. Haider, F., S.d.l. Fuente, and S. Luz, *An Assessment of Paralinguistic Acoustic Features for Detection of Alzheimer's Dementia in Spontaneous Speech*. IEEE Journal of Selected Topics in Signal Processing, 2020. **14**(2): p. 272-281.
31. Ritchie, K., et al., *Recommended cognitive outcomes in preclinical Alzheimer's disease: Consensus statement from the European Prevention of Alzheimer's Dementia project*. *Alzheimers Dement*, 2017. **13**(2): p. 186-195.
32. Vellas, B., et al., *Endpoints for trials in Alzheimer's disease: a European task force consensus*. *Lancet Neurol*, 2008. **7**(5): p. 436-50.
33. Rey, A., *L'examen clinique en psychologie. [The clinical examination in psychology.]*. L'examen clinique en psychologie. 1958, Oxford, England: Presses Universitaires De France. 222-222.
34. Gregory, S., et al., *Involving research participants in a pan-European research initiative: the EPAD participant panel experience*. *Research Involvement and Engagement*, 2020. **6**(1): p. 62.
35. Luz, S. and S. De La Fuente, *A method for analysis of patient speech in dialogue for dementia detection*. , in *Proceedings of LREC 2018*, D. Kokkinakis, Editor. 2018, European Language Resources Association: Paris. p. 7-12.
36. Haider, F., S. De La Fuente, and S. Luz, *An assessment of paralinguistic acoustic features for detection of Alzheimer's Dementia in spontaneous speech*. IEEE Journal of Selected Topics in Signal Processing, 2019. **14**(2): p. 272-281.
37. Gooblar, J., et al., *Attitudes of Research Participants and the General Public Regarding Disclosure of Alzheimer Disease Research Results*. *JAMA neurology*, 2015. **72**(12): p. 1484-1490.
38. Vanderschaeghe, G., et al., *Amnestic MCI patients' experiences after disclosure of their amyloid PET result in a research context*. *Alzheimer's Research & Therapy*, 2017. **9**(1): p. 92.
39. Largent, E.A., et al., *Cognitively unimpaired adults' reactions to disclosure of amyloid PET scan results*. *PLOS ONE*, 2020. **15**(2): p. e0229137.
40. Christensen, K.D., et al., *Disclosing genetic risk for Alzheimer's dementia to individuals with mild cognitive impairment*. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 2020. **6**(1): p. e12002.
41. Grill, J.D., et al., *Short-term Psychological Outcomes of Disclosing Amyloid Imaging Results to Research Participants Who Do Not Have Cognitive Impairment*. *JAMA Neurology*, 2020. **77**(12): p. 1504-1513.

BMJ Open

Remote data collection speech analysis and prediction of the identification of Alzheimer's disease biomarkers in people at risk for Alzheimer's disease dementia and: the Speech on the Phone Assessment (SPeAk) prospective observational study protocol

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2021-052250.R2
Article Type:	Protocol
Date Submitted by the Author:	09-Feb-2022
Complete List of Authors:	Gregory, Sarah; The University of Edinburgh Centre for Clinical Brain Sciences Linz, Nicklas; German Research Centre for Artificial Intelligence Saarbrucken Branch König, Alexandra; National Institute for Research in Computer Science and Automation Langel, Kai; Janssen Healthcare Innovation Pullen, Hannah; The University of Edinburgh Centre for Clinical Brain Sciences Luz, Saturnino; The University of Edinburgh Usher Institute of Population Health Sciences and Informatics Harrison, John; Metis Cognition Ltd; Vrije Universiteit Amsterdam, Department of Neurology Ritchie, Craig; The University of Edinburgh Centre for Clinical Brain Sciences
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Neurology
Keywords:	Dementia < NEUROLOGY, Delirium & cognitive disorders < PSYCHIATRY, Old age psychiatry < PSYCHIATRY

SCHOLARONE™
Manuscripts

1
2
3 Remote data collection speech analysis and prediction of the identification of
4 Alzheimer's disease biomarkers in people at risk for Alzheimer's disease
5 dementia and: the Speech on the Phone Assessment (SPeAk) prospective
6 observational study protocol.
7
8

9 Sarah Gregory¹, Nicklas Linz², Alexandra König³, Kai Langel⁴, Hannah Pullen¹, Saturnino
10 Luz⁵, John Harrison^{6, 7, 8}, Craig Ritchie¹
11
12

13 ¹ Edinburgh Dementia Prevention, Centre for Clinical Brain Sciences, University of
14 Edinburgh, Edinburgh, United Kingdom

15 ² KI Elements, Saarbrücken, Germany

16 ³ National Institute for Research in Computer Science and Automation (INRIA), France

17 ⁴ Janssen Clinical Innovation, Spain

18 ⁵ Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

19 ⁶ Metis Cognition Ltd, Kilmington Common, United Kingdom

20 ⁷ Alzheimer Center Amsterdam, Department of Neurology, Amsterdam UMC, Vrije
21 Universiteit, Amsterdam, The Netherlands

22 ⁸ Institute of Psychiatry, Psychology & Neuroscience, King's College London, United
23 Kingdom
24
25

26
27 Email addresses: Sarah.Gregory@ed.ac.uk; nicklas@ki-elements.de; alexandra@ki-
28 elements.de; KLangel1@its.jnj.com; Hannah.Pullen@ed.ac.uk; s.luz@ed.ac.uk;
29 johnnpc@btinternet.com; craig.ritchie@ed.ac.uk
30
31

32 Corresponding author: Sarah Gregory, Edinburgh Dementia Prevention, Biocube 1, 9 Little
33 France Road, Edinburgh, EH16 4UX; Sarah.Gregory@ed.ac.uk
34

35 Keywords: speech, biomarkers, Alzheimer's disease, cognition
36

37 Word count: 3595 words
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Introduction: Identifying cost effective, non-invasive biomarkers of Alzheimer's disease (AD) is a clinical and research priority. Speech data is easy to collect, and studies suggest it can identify those with AD. We do not know if speech features can predict AD biomarkers in a preclinical population.

Methods and analysis: The Speech on the Phone Assessment (SPeAk) study is a prospective observational study. SPeAk recruits participants aged 50 years and over who have previously completed studies with AD biomarker collection. Participants complete a baseline telephone assessment including spontaneous speech and cognitive tests. A 3-month visit will repeat the cognitive tests with a conversational artificial intelligence bot. Participants complete acceptability questionnaires after each visit. Participants are randomised to receive their cognitive test results either after each visit or only after they have completed the study. We will combine SPeAK data with AD biomarker data collected in a previous study and analyse for correlations between extracted speech features and AD biomarkers. The outcome of this analysis will inform the develop of an algorithm for prediction of AD risk based on speech features.

Ethics and dissemination: This study has been approved by the Edinburgh Medical School Research Ethics Committee (REC reference 20-EMREC-007). All participants will provide informed consent before completing any study related procedures, participants must have capacity to consent to participate in this study. Participants may find the tests, or receiving their scores, causes anxiety or stress. Previous exposure to similar tests may make this more familiar and reduce this anxiety. The study information will include signposting in case of distress. Study results will be disseminated to study participants, presented at conferences and published in a peer reviewed journal. No study participants will be identifiable in the study results.

Strengths and limitations

- A key strength of this study is the novel speech data collection in a preclinical cohort of participants at risk for dementia with AD biomarkers.
- The SPeAk study and associated materials have been designed with input from participants to ensure the project is of interest and value to the target participant population.
- This study will collect data on participant acceptability of the test methodology to collect critical information for future projects.
- Unfortunately in this project there is no concurrent collection of AD biomarkers so only historical data is available.
- The participants recruited to this study do not represent the diversity of the wider population as the previous studies they were involved did not have inclusive recruitment strategies.

Introduction

More than 40 million people worldwide are living with dementia [1]. There is a growing body of work to understand risk factors and how we can prevent dementia [2, 3]. Identification of biomarkers of Alzheimer's disease (AD) has been important in advancing the field [4]. Current gold standard biomarkers for AD include amyloid positron emission tomography (PET) scans [5], cerebrospinal fluid (CSF) amyloid and tau levels [6, 7] and blood based amyloid and tau concentrations [8]. However these techniques can be expensive [9], perceived as invasive or provoking anxiety for some participants [10], and may only be available in certain geographical locations.

Speech data is easy to collect and non-invasive, with recent studies identifying possible markers of AD [11]. Semantic errors and difficulties in semantic processing may be important in preclinical AD [12]. The failure to stop autocorrect errors when reading aloud correlates with A β 1-42 [13]. Changes in connected speech beginning as early as prodromal stages of the illness have been associated with autopsy confirmed AD, alongside linear declines in syntactic complexity and semantic content [14]. Participants with preclinical AD, defined by amyloid beta positivity on a PET scan, were found to have more rapid declines on specific word content, a feature of connected speech [15]. Acoustic features may also reveal information on emotional states and potential presence of neuropsychiatric symptoms such as apathy or depression which are comorbid conditions in AD [16-19].

Adapting traditional face to face cognitive assessments to allow remote data collection may increase access to both research studies and clinical services. Recent studies have shown that automated screening processes, including automated speech analysis, give as accurate neurocognitive screening decisions as human evaluators [20-22]. These tasks are also validated in a telephone-based setting [23], Artificial Intelligence (AI) empowered pipelines can also capture clinically relevant features beyond that available in traditionally administered neurocognitive tests [24, 25].

The aim of the Speech on the Phone Assessment (SPeAk) is to collect speech data from participants across a spectrum of risk for AD, and assess the usefulness of speech features collected over the phone as potential digital biomarkers for AD.

Methods and analysis

Study design

The SPeAk study is a prospective observational study.

Objectives

The objectives of the SPeAk study are:

- To develop algorithms to identify speech biomarkers in the data that are predictive of CSF AD biomarkers.

- To identify ceiling performance on a set of frequently performed cognitive tasks when administered on the telephone.
- To assess the test-retest reliability where the initial testing has a human element and the follow up testing is fully automated
- To explore participant acceptability of both with human and fully automated cognitive testing via the telephone.
- To conduct conversational analysis on the spontaneous speech generated in baseline testing session.

The SPeAK project began in January 2021 and is expected to complete data collection December 2021, with data analysis ongoing until June 2022.

Participants: samples size and power calculation

Participants will all have engaged with in person cognitive testing at the research site within the last 12-24 months. The rationale for this is two-fold. Firstly, some of these participants were enrolled in the European Prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS) [26, 27] and have CSF biomarkers available for analysis in combination with the speech data collected in this protocol. Participants formerly enrolled in the EPAD LCS represent a range of risk states for AD dementia, from cognitively healthy to preclinical AD. Participants recruited from other studies are also either cognitively healthy or have preclinical AD. Importantly none of the participants are currently living with dementia. Secondly, we wanted to recruit participants who were familiar with cognitive testing procedures to ensure they were psychologically robust (as cognitive testing can elicit anxiety) and had a reference point of comparison for acceptance evaluations. All participants included are known to be native or fluent English language speakers from their previous study involvement. If fluent but non-native speakers chose to enroll in the SPeAk study a sensitivity analysis will be conducted to assess for any impact of non-native language patterns.

Participants will be included in this study if they have previously enrolled in the EPAD LCS or another closed commercial research study at the Edinburgh Dementia Prevention research site. Participants must have a landline or mobile phone and have the capacity to consent to the study. Participants will be excluded from the study if they do not have the capacity to provide informed consent or if they anticipate that they will not be available by landline or mobile phone at the three month follow up time point. As participants are being recruited from previous studies we have not specified further eligibility criteria. On the basis of the previous studies all participants will be aged 50 years or above, cognitively healthy or with mild cognitive impairment, will not have dementia, will be fluent English speakers and will not have visual or hearing impairments severe enough to interfere with cognitive testing.

The primary objective is the evaluation of an algorithm, and power calculations remain challenging in this field, with the general accepted wisdom being more data is better. Using the rationale explained by de la Fuente Garcia et al (2019) [28] a minimum of 75 participants would be sufficient for this analysis, however we will aim for at least 150 participants. Briefly in this protocol the authors justify the minimum sample size by placing the lower bounds between $1.2 \times f$ and $1.4 \times f$ where f is the number of features of the dataset. Using the Geneva Minimalistic Acoustic Parameter Set (GeMAPS) [29] which contains 62 features, the authors concluded a minimum of 75 participants would be required for 90% accuracy. In this study we anticipate 30-50 core features. For the machine learning analysis we have adopted a pragmatic approach to sample size calculation, with an aim of 150 participants allowing us to

1
2
3 develop two reasonably sized datasets, one for learning and one for testing. A simple
4 classification algorithm such as the Euclidean distance requires $1.2*v$ where v is the number
5 of variables. If we anticipate 50 core speech features as well as a small number of
6 demographic covariates (described later) we arrive at a required minimum sample size of 66.
7
8

9 **Experimental procedure and design**

10 This study will trial Speech Analysis for cognitive assessments over the telephone in a study
11 population previously exposed to in-person cognitive testing. Participants will be asked to
12 complete two assessments during the study period. In the first testing session participants will
13 interact with a staff member who will deliver the cognitive assessments on the phone,
14 facilitated by an iPad application. The participants do not need to have an iPad themselves,
15 the app will call either a landline or mobile phone number. These cognitive assessments will
16 be audio recorded for speech analysis. There will also be free speech recorded during the
17 telephone conversation which will be used for conversational analysis, and to assess the
18 predictive power and clinical relevance of acoustic features extracted from spontaneous
19 conversational speech [30]. In the follow up testing session, completed 3 months after the
20 first session (+/- 14 days), participants will complete automated cognitive tests. For the
21 automated tests a computerised voice will read the instructions and initiate tasks for the
22 participants. A computer voice was selected instead of a prerecorded human voice to ensure it
23 was clear to participants that they are engaging with a computer and can not freely interact as
24 with a human tester.
25
26
27

28 *Consent*

29 Participants will complete consent forms electronically using Online Survey software
30 (<https://www.onlinesurveys.ac.uk/>). An electronic consent whereby participants will tick
31 consent clauses and type their name has been deemed to be proportionate to the study risks.
32 At the start of the first telephone call for the study, participants will be asked to verbally re-
33 confirm consent to participate in the trial. If the participant does not have the capacity to
34 consent to their participation in the study they will not continue in the study. Members of the
35 research team are experienced in working with this population and assessing capacity. Only
36 appropriately trained staff will be delegated to this task. Senior staff will be available to
37 discuss any concerns about a participant's capacity to consent.
38
39
40
41

42 *Demographics*

43 Demographic details will be collected from all participants after they have provided informed
44 consent to the study. Sex, age, education (years), living status and current medications will be
45 recorded.
46
47

48 *Preparing the participant*

49 Prior to the spontaneous speech and cognitive tests the participants will be prepared for the
50 session by the interviewer. The interviewer will check the audio quality, both for the
51 participant and for the interviewer. The participants will be asked to position themselves in a
52 quiet room away from distractions and to ensure they do not have anything available to write
53 on (such as pen and paper or an electronic device). Participants will be informed that if the
54 call is terminated for any reason an attempt will be made to re-establish the call. Prior to the
55 second call, which is led by the computer, the participants will be reminded of these
56 preparations in an email. If participants would like a phone call to review these preparations
57 this will be offered.
58
59
60

Spontaneous speech

Participants will be informed the rest of the session will be audio recorded and the session will begin with approximately 5 minutes of spontaneous speech, prompted with the completion of two conversational tasks. In the first task participants will be asked to respond to the question: 'What animal makes the best pet?'. To encourage conversational exchange between the interviewer and the participant, the interviewer will ask follow-up questions, such as 'Why do you think that?' to elicit turn-taking. In task two participants will be invited to engage in a game of 20 questions. The interviewer will explain briefly that it is a game in which one participant thinks of an object, animal, plant or substance and the other participant asks yes-or-no questions to determine what it is. To offer a level of standardisation the interviewer will always think of the same animal and the participant will ask questions to try and guess the answer. The participant may only ask questions that can be answered with a yes or a no. This data will only be collected at the baseline visit as a human interviewer is needed for this.

Cognitive testing

Participants will complete a series of cognitive tasks selected to test a range of domains and assessed for suitability to complete on the telephone. Domains were selected to represent cognitive abilities sensitive to decline in early AD [31, 32]. Encoding, immediate and delayed recall of verbal words will be tested using the immediate and delayed Rey Auditory Visual Learning Test (RAVLT) list learning task [33]. Between the initial and delayed recall tasks participants will also be asked to complete the Wechsler Adult Intelligence Scale-IV (WAIS-IV) forward digit span, phonemic (letter S) and semantic (category animals) verbal fluency tasks. All test instructions will be read verbatim by a human interviewer at baseline and a computerised voice at 3-month follow up. An audio recording is made of participant responses which will be used for automatic transcription and scoring of the tasks.

Results feedback

Participants will be randomised 1:1 to receive results of their cognitive tests after each session or only after the final session. This is to analyse whether participants have a preference for receiving immediate feedback or are happy to receive feedback at the end of their involvement in the study. No individual level speech feature data will be feedback to participants as this remains experimental. Participants will be provided with a copy of the overall study results. Individual cognitive test results will be sent to the participant using a proforma that explains the experimental nature of the tasks (i.e. being completed on the phone rather than face to face), the expected score range for their age (based on published normative range for face to face assessment) and their scores. This proforma will be sent via email from the research team. If a participant has a score that falls below the published normative range their data will be reviewed by the Chief Investigator (CWR, an experienced Consultant Psychiatrist, clinical triallist and Professor of Old Age Psychiatry) prior to result feedback. Any results that may be of concern will be highlighted to the participant in a phone call prior to receiving the email with their results and a letter to their GP written if the participant consents to this.

Acceptability questionnaires

Following each assessment point, participants will complete an acceptability questionnaire designed for this study. These questionnaires will evaluate how easy the assessments were to set up, how comfortable participants felt completing these assessments remotely, whether participants preferred face to face assessments (completed in earlier trials), human tester

1
2
3 administered assessments or computer administered assessments and understand participants
4 experiences of receiving their test results.
5

6 *Previously acquired biomarker data*

7
8 Participants have previously taken part in studies which collected AD biomarkers (CSF, PET
9 and structural MRI), information about AD risk factors (such as apolipoprotein E genetic
10 status) and cognitive test batteries (Repeatable Battery for the Assessment of
11 Neuropsychological Status (RBANS), Clinical Dementia Rating scale (CDR), Mini Mental
12 State Examination (MMSE)). These data will be used in the machine learning model to build
13 the algorithm for the primary outcome.
14
15

16 **Data Management**

17
18 All speech data will be recorded on the Delta Testing App on an iPad and securely
19 transferred to a University of Edinburgh server. Participants are all assigned a unique ID code
20 which is used to label all of their study data. Due to the potential benefit of speech data for
21 future research, participants will be asked to consent to long term storage of their raw speech
22 files. These files will be securely held on University of Edinburgh servers and will be
23 available for analysis for the research team as part of this protocol, and in the future for
24 approved studies from researchers approved by the study team. Data from the EPAD LCS
25 study are held securely on the Alzheimer's Drug Discovery Foundation (ADDF) platform
26 and relevant participant data will be extracted for use in planned analyses. Demographic and
27 questionnaire data will initially be stored using onlinesurveys.ac.uk before transfer to the
28 study server. All procedures are in line with the University of Edinburgh data protection
29 policy which is informed by the General Data Protection Regulation (GDPR).
30
31
32

33 **Patient and public involvement statement**

34 Two participants from the EPAD LCS participant panel [34] were involved in the study
35 design process and reviewed the study documentation electronically. In particular the
36 participants were asked their opinions on whether they thought it was acceptable to conduct
37 cognitive assessments on the telephone and their thoughts on receiving results from the
38 assessments. Both participants agreed that completing cognitive assessments on the telephone
39 would be acceptable to them, and they felt that would apply to the rest of the group. Both
40 welcomed the feedback of results as this was something the panel had advocated for with
41 previous studies. The participants felt that it should be clear in the information sheet that
42 feedback would be available, so that if anyone who did not want to receive this would know
43 to decline the study invitation. All comments on the participant facing documentation were
44 incorporated into the final document to improve readability and maximise ease of use and
45 understanding of the study documents.
46
47
48

49 **Analysis**

50 To achieve the primary objective of the study, three different types of variables will be
51 computed and extracted from speech recordings: classical neuropsychological outcome
52 variables, novel or qualitative outcome variables based on produced language and low-level
53 speech descriptors. These measures will be computed on a task-level, meaning for each
54 cognitive task performed. Language and speech variables will be extracted on a linguistic and
55 acoustic level based. These include temporal, voice source, formant, semantic, and syntactic
56 variables. Some variables are task specific and encode performance or strategy related to a
57 task, some are general descriptors of voice quality. For the primary analysis we will develop
58 machine learning models, with patients' biomarker status as a target variable. Models will be
59
60

1
2
3 constructed with task specific features and using aggregated variables spanning multiple
4 cognitive tasks. Patients with missing data will not be considered in the final analysis. An
5 interim analysis will be performed once 100 participants have been recorded.
6 Secondary objectives are to evaluate ceiling effects and stability in cognitive performance
7 between a human and a computer tester. For this analysis classical neuropsychological outcome
8 variables will be extracted from the recorded audio. To this end, human raters will use classical
9 scoring schemes to score speech recordings of cognitive tasks. This will lead to a dataset where
10 each participant has multiple repeated measurements of the same task. Results from multiple
11 time points will be compared using a repeated measures ANOVA and repeated measures
12 correlation. Patients with missing data will not be considered in the final analysis. An interim
13 analysis will be performed once 100 participants have been recorded. We will use frequency
14 statistics and qualitative analysis for the acceptability questionnaires and compare baseline and
15 3-month responses used paired t-tests. We will also analyse and predict levels of participant
16 satisfaction, as recorded in the acceptability questionnaires, based on acoustic and
17 conversational speech features.
18
19
20

21 Analysis of the conversational (dialogue) data will generate speech interaction features and
22 vocalisation graphs which will be used as input representations for machine learning
23 modelling. These interactional features will be combined with low level acoustic descriptors
24 for predictive modelling of biomarker status and results of the cognitive tasks [35, 36].
25
26

27 **Ethics and dissemination**

28 This study has been reviewed and given favourable ethical opinion by the Edinburgh Medical
29 School Research Ethics Committee (EMREC) (REC Reference 20-EMREC-007). The study
30 raises few ethical concerns. Participants may find participation in cognitive testing causes
31 stress or anxiety. All participants invited to this study will previously have completed at least
32 one testing session in a face-to-face setting. We hope that this will reduce the risk of
33 participants experiencing stress as the cognitive tasks will be somewhat routine and
34 participants who did find the cognitive testing anxiety provoking in the preceding studies are
35 likely to decline this study invitation. Participants with concerns about their memory will be
36 encouraged to speak to their primary care practitioner or local Alzheimer's charity hotlines.
37
38
39

40 All participants will provide informed consent prior to participating in the SPeAk study. We
41 will use electronic consent for this study, via the Online Survey software
42 (<https://www.onlinesurveys.ac.uk/>). This software provides an electronic consent whereby
43 participants will tick consent clauses and type their name. Electronic consent has been
44 deemed to be proportionate to the SPeAk study risks. At the start of the first telephone call
45 for the study, participants will be asked to verbally re-confirm consent to participate in the
46 trial. If the participant does not have the capacity to consent to their participation in the
47 study they will not continue in the study. Members of the research team are experienced in working
48 with this population and assessing capacity. Only appropriately trained staff will be delegated
49 to this task. Senior staff will be available to discuss any concerns about a participant's
50 capacity to consent.
51
52
53

54 Receiving results from the cognitive tests may also lead participants to worry about their
55 memory. Providing participants with their results is increasingly common, and the plan to do
56 so in this study was well received by our participant panel advisors. Nearly half of
57 participants enrolled in a longitudinal AD study were interested to receive their cognitive test
58 results [37]. In that study participants highlighted receiving personal feedback would be a
59 motivation for engagement in longitudinal research studies [37]. We were not able to identify
60

1
2
3 studies reporting on participants' experiences of receiving their cognitive test results and will
4 evaluate this within the acceptability questionnaires in our study. Studies that have
5 investigated disclosure of AD risk factors, such as amyloid PET and APOE results, have
6 identified both benefits and harms of disclosure [38-40]. When done in a safe and appropriate
7 way, there is little short term psychological harm in risk factor disclosure [41]. On balance
8 disclosure, when done in a safe and sensible way, allows participants to plan and prepare
9 causing little psychological distress. As previously detailed, any participants who express
10 particular concern about their cognitive test scores will be encouraged to seek help from their
11 primary care practitioner or Alzheimer's disease charities.
12
13

14
15 The final ethical consideration is data security. Participants voices will be recorded, stored
16 and analysed. The study has been reviewed in line with university information governance
17 guidelines and complies with data security requirements. Publications arising from this study
18 will not identify any participants. Participants will be asked to explicitly consent to the
19 recording, storage and analysis of their voice data. Any participants who do not agree to this
20 will not be enrolled in the study.
21
22

23 **Discussion**

24 The SPeAk study adds a unique opportunity to identify speech biomarkers in a preclinical
25 population at risk for dementia. The study will combine the collection of spontaneous speech
26 through two conversational tasks and a brief battery of cognitive tasks previously validated
27 in-person. The primary aim of the study is to develop an algorithm to identify speech features
28 which are predictive of the target variable of AD biomarkers (collected in a prior study).
29 Secondary objectives will evaluate test-retest reliability of the cognitive tasks when delivered
30 on the phone and participant acceptability of the assessments, including experiences of
31 receiving their cognitive test results. We will also undertake exploratory conversational
32 analysis of the spontaneous speech data.
33
34

35 Early detection of Alzheimer's disease is a priority for both clinical and research purposes.
36 Identifying a low-cost, non-invasive biomarker will be important in this field, and speech is
37 an exciting biomarker to explore.
38
39

40 **Acknowledgments**

41 We would like to thank the EPAD Scotland participant's panel for their review of the study,
42 and for all prospective participants for taking the time to consider this study.
43
44

45 **Funding**

46 This work is supported by Janssen Pharmaceutica NV through a collaboration agreement
47 (Award/Grant number is not applicable).
48
49

50 **Competing interests**

51 KL is an employee of Janssen (the funder). NL & AK are employees of KI Elements who
52 own and provide the Delta Testing App. JH reports receipt of personal fees in the past 2 years
53 from Actinogen, AlzeCure, Aptinyx, Astra Zeneca, Athira Therapeutics, Axon Neuroscience,
54 Axovant, Bial Biotech, Biogen Idec, BlackThornRx, Boehringer Ingelheim, Brands2life,
55 Cerecin, Cognito, Cognition Therapeutics, Compass Pathways, Corlieve, Curasen, EIP
56 Pharma, Eisai, G4X Discovery, GfHEU, Heptares, Ki Elements, Lundbeck, Lysosome
57 Therapeutics, MyCognition, Neurocentria, Neurocog, Neurodyn Inc, Neurotrack, the NHS,
58 Novartis, Novo Nordisk, Nutricia, Probiobdrug, Prothena, Recognify, Regeneron, reMYND,
59 Rodin Therapeutics, Samumed, Sanofi, Signant, Syndesi Therapeutics, Takeda, Vivoryon
60

1
2
3 Therapeutics and Winterlight Labs. Additionally, he holds stock options in Neurotrack Inc.
4 and is a joint holder of patents with My Cognition Ltd. CWR has received consultancy fees
5 from Biogen, Eisai, MSD, Actinogen, Roche, and Eli Lilly, as well as payment or honoraria
6 from Roche and Eisai. SG, HP & SL have no conflicts to declare.
7
8

9 **Author Statement**

10 SG coordinated the design of the protocol, managed the ethical approval process, drafted the
11 manuscript and edited the final version. NL, AK, KL, HP, SL and JH all contributed to the
12 design of the protocol and provided comments on the manuscript drafts. CWR is the principal
13 investigator of the protocol and oversaw the design and ethical approval process as well as
14 provided comments on the manuscript drafts.
15
16

17 **Data statement**

18 Data will be available through a data access process from the University of Edinburgh. This
19 data access will be managed by the study Principal Investigator (CWR), study coordinator
20 (SG) in conjunction with the University of Edinburgh data management specialists. As
21 speech data is potentially identifiable it is not appropriate for this data to be submitted to
22 open data access platforms external to the University of Edinburgh.
23
24
25

26 **References**

- 27 1. Nichols, E., et al., *Global, regional, and national burden of Alzheimer's disease and*
28 *other dementias, 1990–2016: a systematic analysis for the Global Burden of*
29 *Disease Study 2016*. The Lancet Neurology, 2019. **18**(1): p. 88-106.
- 30 2. Livingston, G., et al., *Dementia prevention, intervention, and care: 2020 report of the*
31 *Lancet Commission*. The Lancet, 2020. **396**(10248): p. 413-446.
- 32 3. Livingston, G., et al., *Dementia prevention, intervention, and care*. The Lancet, 2017.
33 **390**(10113): p. 2673-2734.
- 34 4. Humpel, C., *Identifying and validating biomarkers for Alzheimer's disease*. Trends in
35 biotechnology, 2011. **29**(1): p. 26-32.
- 36 5. Chételat, G., et al., *Amyloid-PET and ¹⁸F-FDG-PET in the diagnostic*
37 *investigation of Alzheimer's disease and other dementias*. The Lancet Neurology,
38 2020. **19**(11): p. 951-962.
- 39 6. Blennow, K., et al., *Predicting clinical decline and conversion to Alzheimer's disease*
40 *or dementia using novel Elecsys Aβ(1–42), pTau and tTau CSF immunoassays*.
41 Scientific Reports, 2019. **9**(1): p. 19024.
- 42 7. Blennow, K. and H. Zetterberg, *Biomarkers for Alzheimer's disease: current status*
43 *and prospects for the future*. Journal of Internal Medicine, 2018. **284**(6): p. 643-663.
- 44 8. Hampel, H., et al., *Blood-based biomarkers for Alzheimer disease: mapping the road*
45 *to the clinic*. Nature Reviews Neurology, 2018. **14**(11): p. 639-652.
- 46 9. Wittenberg, R., et al., *Economic impacts of introducing diagnostics for mild cognitive*
47 *impairment Alzheimer's disease patients*. Alzheimer's & dementia (New York, N. Y.),
48 2019. **5**: p. 382-387.
- 49 10. Peskind, E.R., et al., *Safety and acceptability of the research lumbar puncture*.
50 Alzheimer Dis Assoc Disord, 2005. **19**(4): p. 220-5.
- 51 11. Boschi, V., et al., *Connected Speech in Neurodegenerative Language Disorders: A*
52 *Review*. Front Psychol, 2017. **8**: p. 269.
53
54
55
56
57
58
59
60

12. Venneri, A., et al., *Diagnostic and prognostic role of semantic processing in preclinical Alzheimer's disease*. *Biomarkers in Medicine*, 2018. **12**(6): p. 637-651.
13. Gollan, T.H., et al., *Failure to stop autocorrect errors in reading aloud increases in aging especially with a positive biomarker for Alzheimer's disease*. *Psychology and Aging*, 2020. **35**(7): p. 1016-1025.
14. Ahmed, S., et al., *Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease*. *Brain*, 2013. **136**(12): p. 3727-3737.
15. Mueller, K.D., et al., *Amyloid beta associations with connected speech in cognitively unimpaired adults*. *Alzheimers Dement (Amst)*, 2021. **13**(1): p. e12203.
16. Trevino, A.C., T.F. Quatieri, and N. Malyska, *Phonologically-based biomarkers for major depressive disorder*. *EURASIP Journal on Advances in Signal Processing*, 2011. **2011**(1): p. 42.
17. Yang, Y., C. Fairbairn, and J.F. Cohn, *Detecting Depression Severity from Vocal Prosody*. *IEEE transactions on affective computing*, 2013. **4**(2): p. 142-150.
18. Low, D.M., K.H. Bentley, and S.S. Ghosh, *Automated assessment of psychiatric disorders using speech: A systematic review*. *Laryngoscope Investigative Otolaryngology*, 2020. **5**(1): p. 96-116.
19. König, A., et al., *Detecting Apathy in Older Adults with Cognitive Disorders Using Automatic Speech Analysis*. *J Alzheimers Dis*, 2019. **69**(4): p. 1183-1193.
20. König, A., et al., *Use of Speech Analyses within a Mobile Application for the Assessment of Cognitive Impairment in Elderly People*. *Curr Alzheimer Res*, 2018. **15**(2): p. 120-129.
21. König, A., et al., *Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task*. *Dement Geriatr Cogn Disord*, 2018. **45**(3-4): p. 198-209.
22. Themistocleous, C., M. Eckerström, and D. Kokkinakis, *Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls*. *PLOS ONE*, 2020. **15**(7): p. e0236009.
23. Tröger, J., et al., *Telephone-based Dementia Screening I: Automated Semantic Verbal Fluency Assessment*, in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2018, Association for Computing Machinery: New York, NY, USA. p. 59–66.
24. Linz, N., et al., *Using Neural Word Embeddings in the Analysis of the Clinical Semantic Verbal Fluency Task*. *IWCS 2017 — 12th International Conference on Computational Semantics — Short papers*, 2017.
25. Linz, N., et al. *Predicting Dementia Screening and Staging Scores from Semantic Verbal Fluency Performance*. in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. 2017.
26. Ritchie, C.W., et al., *The European Prevention of Alzheimer's Dementia (EPAD) Longitudinal Cohort Study: Baseline Data Release V500.0*. *The Journal of Prevention of Alzheimer's Disease*, 2020. **7**(1): p. 8-13.
27. Ritchie, C.W., et al., *Development of interventions for the secondary prevention of Alzheimer's dementia: the European Prevention of Alzheimer's Dementia (EPAD) project*. *The Lancet Psychiatry*, 2016. **3**(2): p. 179-186.
28. de la Fuente Garcia, S., C.W. Ritchie, and S. Luz, *Protocol for a conversation-based analysis study: PREVENT-ED investigates dialogue features that may help predict dementia onset in later life*. *BMJ Open*, 2019. **9**(3): p. e026254.

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
29. Eyben, F., et al., *The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing*. IEEE Transactions on Affective Computing, 2016. **7**(2): p. 190-202.
30. Haider, F., S.d.l. Fuente, and S. Luz, *An Assessment of Paralinguistic Acoustic Features for Detection of Alzheimer's Dementia in Spontaneous Speech*. IEEE Journal of Selected Topics in Signal Processing, 2020. **14**(2): p. 272-281.
31. Ritchie, K., et al., *Recommended cognitive outcomes in preclinical Alzheimer's disease: Consensus statement from the European Prevention of Alzheimer's Dementia project*. *Alzheimers Dement*, 2017. **13**(2): p. 186-195.
32. Vellas, B., et al., *Endpoints for trials in Alzheimer's disease: a European task force consensus*. *Lancet Neurol*, 2008. **7**(5): p. 436-50.
33. Rey, A., *L'examen clinique en psychologie. [The clinical examination in psychology.]*. L'examen clinique en psychologie. 1958, Oxford, England: Presses Universitaires De France. 222-222.
34. Gregory, S., et al., *Involving research participants in a pan-European research initiative: the EPAD participant panel experience*. *Research Involvement and Engagement*, 2020. **6**(1): p. 62.
35. Luz, S. and S. De La Fuente, *A method for analysis of patient speech in dialogue for dementia detection*. , in *Proceedings of LREC 2018*, D. Kokkinakis, Editor. 2018, European Language Resources Association: Paris. p. 7-12.
36. Haider, F., S. De La Fuente, and S. Luz, *An assessment of paralinguistic acoustic features for detection of Alzheimer's Dementia in spontaneous speech*. IEEE Journal of Selected Topics in Signal Processing, 2019. **14**(2): p. 272-281.
37. Gooblar, J., et al., *Attitudes of Research Participants and the General Public Regarding Disclosure of Alzheimer Disease Research Results*. *JAMA neurology*, 2015. **72**(12): p. 1484-1490.
38. Vanderschaeghe, G., et al., *Amnestic MCI patients' experiences after disclosure of their amyloid PET result in a research context*. *Alzheimer's Research & Therapy*, 2017. **9**(1): p. 92.
39. Largent, E.A., et al., *Cognitively unimpaired adults' reactions to disclosure of amyloid PET scan results*. *PLOS ONE*, 2020. **15**(2): p. e0229137.
40. Christensen, K.D., et al., *Disclosing genetic risk for Alzheimer's dementia to individuals with mild cognitive impairment*. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 2020. **6**(1): p. e12002.
41. Grill, J.D., et al., *Short-term Psychological Outcomes of Disclosing Amyloid Imaging Results to Research Participants Who Do Not Have Cognitive Impairment*. *JAMA Neurology*, 2020. **77**(12): p. 1504-1513.