

REVIEW ARTICLE

The collection of speech data for the assessment of cognition remotely: Balancing ethical and practical challenges

Stina Saunders¹  | Fasih Haider¹ | Alyssa M. Lanzi² | Brian MacWhinney³ |
Davida Fromm³ | Chia-Ju Chou⁴ | Yi-Chien Liu⁴ | Ya-Ning Chang⁵ |
Sarah Gregory^{6,7} | Craig Ritchie⁷ | Ioulietta Lazarou⁸ | Spiros Nikolopoulos⁸ |
Ioannis Kompatsiaris⁸ | Samuel O. Danso^{9,10} | Graciela Muniz-Terrera^{6,11} |
Juan Rafael Orozco-Arroyave^{12,13} | Paula Andrea Perez-Toro^{12,13} |
Raquel Rangel Cesario^{14,15} | Manuel Cesario¹⁴ | Saturnino Luz¹ 

¹Usher Institute, School of Population Health Sciences, The University of Edinburgh, Edinburgh, UK

²Department of Communication Sciences and Disorders, University of Delaware, Newark, Delaware, USA

³Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

⁴Department of Neurology, Cardinal Tien Hospital, New Taipei City, Taiwan

⁵Miin Wu School of Computing, National Cheng Kung University, Tainan City, Taiwan

⁶Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh BioQuarter, Edinburgh, UK

⁷School of Medicine, University of St. Andrews, St. Andrews, UK

⁸Centre for Research and Technology Hellas, Information Technologies Institute, Thessaloniki, Greece

⁹School of Computer Science and Engineering, University of Sunderland, Sunderland, UK

¹⁰School of Allied Health Sciences, University of Cape Coast, Cape Coast, Ghana

¹¹Heritage College of Osteopathic Medicine, Ohio University, Athens, Ohio, USA

¹²GITA Lab, Universidad de Antioquia, Antioquia, Colombia

¹³Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

¹⁴Magdala Academy, São Paulo, Brazil

¹⁵Medicina, Uni-FACEF, São Paulo, Brazil

Correspondence

Stina Saunders and Saturnino Luz, Usher Institute, School of Population Health Sciences, College of Medicine and Veterinary Medicine, The University of Edinburgh, 5-7 Little France Rd, Edinburgh EH16 4UX, UK. Email: stina.saunders@ed.ac.uk and s.luz@ed.ac.uk

Funding information

UKRI, Grant/Award Number: 10102226

Abstract

Speech biomarkers could form a critical step in improving the accessibility, scalability, and early detection of Alzheimer's disease and related dementias. However, ethical and practical challenges remain across regulatory and cultural contexts. In this paper, we briefly review the challenges in adopting speech biomarkers, relate our experiences globally to recent advances in the neurodegenerative field, and consider how speech assessments could be integrated into clinical care. Insights from high- and low- and middle-income countries (Taiwan, Ghana, Colombia, Brazil, Greece, UK, and US) demonstrate the potential impact of speech technology. While there are common benefits, risks, incentives, and hurdles, many aspects are specific to country or

Stina Saunders and Saturnino Luz are joint-first authors, and contributed equally to the paper.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

region. There is a need for more speech data sets (particularly in languages other than English), standardization in data collection and analysis, stronger collaboration between machine learning and neurodegenerative disease experts, unified privacy regulation, and finally, a consensus on the clinical interpretation of speech biomarker data.

KEYWORDS

Alzheimer's disease, brain health, detection of cognitive decline, outcome measures, speech biomarkers

Highlights

- Remote speech-based biomarkers have potential as scalable tools for early neurodegenerative disease detection across varied health-care systems globally.
- We present insights from seven countries highlighting both shared benefits and region-specific hurdles to implementation and equity of speech-based biomarkers.
- Clinical adoption depends on validation of norms, clinician training, and culturally adapted tools.
- Global disparities in available large language models and privacy regulation hinder development and deployment of speech biomarkers.
- Remote assessments can increase clinical reach in rural populations, supporting more inclusive brain health care worldwide.

1 | BACKGROUND

Neurodegenerative brain diseases impact areas of speech long before overt clinical symptoms of Alzheimer's disease (AD) and related dementias (ADRD) appear, influencing aspects such as richness of vocabulary, fluency, and acoustic features of speech.¹ Consequently, speech-based biomarkers have been proposed as a cost-effective method for the early identification, prognosis, and management of ADRD.² The primary appeal of speech biomarkers lies in their ease of use for seldom heard populations and longitudinal monitoring across diverse cohorts and diseases.^{3,4}

Recent research has accelerated speech-based biomarkers as indicators of preclinical, prodromal, and symptomatic neurodegenerative disease. Research in this area is increasingly shifting from single-feature acoustic analyses toward integrative models incorporating both linguistic and paralinguistic features, reflecting the fact that neurodegeneration alters multiple layers of speech production.⁵ While cognitive tests administered in clinical environments have included speech tasks such as category naming, text reading, and picture description, speech data collection has become more feasible in naturalistic settings. In one recent multi-day remote tablet-based study of cognitively unimpaired individuals, for instance, > 90% of participants adhered to the study protocol, and averaged acoustic features showed good test-retest reliability and high usability scores.⁶ The same study observed differences in specific acoustic markers (e.g., pause-to-word

ratio) between amyloid beta ($A\beta$)-positive and -negative participants, although the higher pause-to-word ratio in $A\beta$ -positive participants did not remain significant after correction for multiple comparisons.

Similarly, Cho et al.⁷ developed an automated language processing framework that distinguished mild cognitive impairment (MCI) from no impairment with high accuracy using naturalistic narrative speech. Their approach uses transformer-based language models trained on age-matched data, to detect subtle lexical-semantic shifts indicative of early cognitive decline ([Supporting Information](#)).

This growing body of evidence supports the potential clinical relevance of speech biomarkers for early-stage AD, though many key aspects of their reliable use remain challenging, including determining the etiology of decline. This paper discusses the utility of speech biomarkers in the detection, monitoring, and management of clinical syndromes associated with ADRD in a global context, and explores the value, risks, and future applications of using speech biomarkers to remotely monitor cognition and brain health.

2 | THE MAIN ADVANTAGES AND DISADVANTAGES OF SPEECH BIOMARKERS

Assessment of neurodegenerative disease typically requires testing in person, conducted by trained and qualified team members. In contrast, remote assessments offer a valuable alternative for both assessment

and ongoing care in ADRD,⁸ reducing travel burden for the person and care partner and the need for specialist staff time. Additionally, speech biomarkers and telemedicine widen accessibility to seldom heard groups and increase scalability of both early detection and monitoring of cognitive function. For instance, the DementiaBank protocol in the United States was specifically designed to be delivered via telehealth and enables reaching participants in remote rural locations as well as population groups who are typically underrepresented in research,⁹ although noise-reduction and compression algorithms associated with different platforms may impact the types of speech acoustics analyses that can be validly and reliably extracted. Moreover, the use of everyday digital devices in remote monitoring enables frequent, ecologically valid data collection.

The argument in favor of speech-based biomarkers in AD lies in their sensitivity to early cognitive changes as well as the possibility of detecting and monitoring wider well-being, such as mood disorders¹⁰ and apathy,¹¹ which often overlap with AD. Speech biomarkers are an area with rapidly growing evidence around their utility in both research and clinical contexts,¹² demonstrating a demand for remote, scalable assessment methods in neurodegenerative disease research and clinical practice. The remote capability is especially valuable in countries where there are difficult-to-access regions, but also ongoing internal conflict where access is restricted by guerrilla activity. Additionally, in countries like Ghana, where there are strong cultural beliefs against blood-based approaches to screening,¹³ non-invasive speech-based biomarkers present a possibly more appropriate and acceptable alternative for early detection and monitoring of cognitive decline. Further, in Ghana, Colombia, Brazil, and many other low- and middle-income countries (LMICs), remote data collection and telehealth are becoming increasingly feasible and commonly used.¹⁴ This potential for broadening access aligns with the targets of the World Health Organization (WHO)'s global action plan for public health responses to dementia, particularly in regard to risk reduction and diagnosis, treatment, and care, including the target that at least 50% of countries achieve a diagnosis rate of at least 50%, which remains far from being met.¹⁵

Finally, speech-based assessments that predict amyloid status^{6,16,17} would allow non-invasive screening of AD pathology and eligibility assessment for anti-amyloid antibody treatments.¹⁸ On these grounds, speech biomarkers appear to be a critical component in addressing the multifaceted challenges in managing ADRD.

Despite these advantages, there are several concerns associated with speech biomarkers, including cultural insensitivity,¹⁹ demographic biases,²⁰ privacy concerns,²¹ and the lack of interdisciplinary collaboration between experts in statistical methods, machine learning/artificial intelligence (AI), and clinical neuroscience. Furthermore, there is no consensus around how to standardize data acquisition protocols, such as the volume and variety of speech data needed to ensure the accurate tracking of pathological changes.¹²

With more than half of the world's population speaking more than one language,²² it is critical to derive speech norms for non-European languages²³ as well as adapt and update language technology components and models that apply in bilingual and multilingual settings, including changes in language over time. While new technologies have

enabled data collection on relatively large scales, there are still several technological challenges, such as accounting for bilingualism, in which frequent language switching in multilingual households complicates data collection and modeling. While some of these challenges could be tackled, in principle, by incorporating such covariates into predictive models, doing so requires careful design.

The main limitation in studies analyzing languages other than English is that most language models, notably large language models (LLMs), have been trained mostly on English speech data, and very few LLMs trained primarily in other languages exist.²⁴ Other kinds of machine learning and signal processing models used in typical speech processing systems, such as acoustic feature extractors, automatic speech recognition, parsers, and so forth, also struggle to cope with low-resource languages. This is also true of foundation models that target acoustic data,²⁵ for which specific sounds of certain languages are often not adequately modeled due to data scarcity. For instance, our experience in Taiwan illustrates this issue: we have observed that machine learning models trained on limited datasets in Mandarin Chinese language may yield a high rate of false positive cases when applied to broader populations, risking clinical mismanagement of these false positive cases. Furthermore, for languages such as Taiwanese or Twi (spoken in Ghana) computational methods for automatic speech recognition and language analysis are incompletely developed. Consequently, specific lexical units and morphosyntax phenomena in these linguistic contexts are not appropriately understood in the context of ADRD.

3 | PROTECTING PERSONALLY IDENTIFIABLE INFORMATION AND PRIVACY

It is recognized that an audio recording may carry personal information that the speaker may not intend or expect to disclose.²⁶ In the context of ADRD, speech data may incorporate acoustic speech features like pitch (how high or low a speaker's voice sounds), jitter (inconsistencies or wobbles in the voice), articulation rate (speed of speaking), phonemic identifiability (the probability of a certain phoneme being correctly pronounced), and semantic coherence (words making sense),^{12,27} making speech data personally identifiable information, with concerns around data storage, sharing, and the risk of re-identification.³

Unlike structured numeric data, making speech data anonymous can be complex and resource-intensive, often requiring machine learning expertise, staff time, and appropriate programs. Some techniques used to protect the privacy of speech data include: (1) encrypting audio files for storage and transfer; (2) splitting data into random components, each processed independently to reduce the risk of privacy leakage; (3) learning data representations, such as through auto-encoder technologies that remove identifiable information;^{28,29} (4) transfer learning, which consists of creating models that incorporate information from different locations where data collection was performed, that is, only transferring locally generated parameters without sharing any recording among different hospitals or research centers;³⁰ (5) voice transformation to alter the person's acoustic features while preserving

clinically relevant speech characteristics; and (6) differential privacy, which introduces statistical noise to the dataset.³¹

Although privacy concerns can be addressed to a certain extent with the use of privacy-preserving machine learning techniques, the broader challenge lies in the limited regulation governing large-scale data acquisition. For instance, current privacy legislation, which is limited to more traditional data protection principles such as minimizing the amount of data collected and informed consent, based on the global Fair Information Practices (FIPs), may not sufficiently consider the depth of information that can be derived from speech data.³²

There is no global standard for speech data governance, and country- and region-specific regulations vary considerably. In Latin American countries, while privacy regulations exist,³³ they are generally more lenient than those in the EU. Furthermore, the application of regulations can vary significantly across different nations and specific contexts, particularly when dealing with sensitive information such as genetic data. In Brazil, the data protection legislation (“Lei geral de proteção de dados” [LGPD]) was introduced in 2018 and regulates rights and access to sensitive information. There has been vigorous debate since then on the implications of the LGPD to AI research, particularly in health technology.³⁴ In Colombia, which hosts one of the largest populations with genetic early onset AD,³⁵ privacy concerns are particularly acute. Although initiatives like RedLat³⁶ and BrainLat³⁷ have advanced collaborative dementia research across Latin America, data protection remains a pressing concern in cross-border collaborations due to the need to navigate individual countries’ privacy regulations.

A proposed solution is the increasing use of trusted research environments (TRE), that is, platforms that both store and enable analysis of research data, negating the need for data to be transferred out of the secure platform and allowing access to users within the same platform.³⁸ Examples include the Alzheimer’s Disease Data Initiative’s (ADDI) online platform (<https://www.alzheimersdata.org/>), the AD Workbench (<https://www.alzheimersdata.org/ad-workbench>) and DATAMIND (<https://datamind.org.uk>).

4 | LEGAL, INSURANCE, AND CAPACITY CONCERNS IN LONGITUDINAL MONITORING

The collection and longitudinal use of speech data in neurodegenerative disease present complex issues around data privacy and anonymity, consent in how speech data are used, and who the data are shared with. In the EU and UK, frameworks such as the General Data Protection Regulation (GDPR) necessitate strict regulatory compliance, including clear consent for speech data collection. In Brazil, the LGPD defines a national framework similar to the GDPR. While there is no federal law like the GDPR in the United States, personal information handled in connection with speech and language services conducted in health-care settings is covered by the Health Insurance Portability and Accountability Act (HIPAA) as well as state laws around data protection. This situation is similar in countries like Colombia, where the regulation is not clear and most responsibility for patients’ privacy and data rests on the Research Ethics Committees’ criteria.

While there is no current precedent for insurance companies to access speech data, evidence around individuals’ preferences in sharing other types of AD biomarker results suggests there are concerns around stigma and discrimination with regard to sharing AD assessment results.³⁹ It remains a gray area how much individuals would be expected or legally obliged to disclose to insurance companies regarding speech-based findings on health status, and how, if at all, insurance companies would be able to interpret results from novel AD biomarker assessments.

There is no consensus in the literature on how to involve participants in longitudinal speech assessments when cognitive decline may impact capacity over time, raising ethical questions about consent and ongoing participation in the AD population. In systems such as those in Scotland, ascertaining and recording ongoing assent at each study visit aligns with both the Adults with Incapacity (Scotland) Act, 2000⁴⁰ and Good Clinical Practice guidelines⁴¹ even after capacity diminishes. However, enrolling individuals in research when there is diminished capacity to consent even at study entry can pose challenges, particularly when communicating about complex data protection issues that should be considered prior to study entry. Designing clear communication aids around what speech data will be collected, for what purpose, and what safeguards are in place, to support participant and care partner decision making are key. Furthermore, partnering with people with a lived experience to design the study and participant-facing materials can be hugely beneficial when dealing with digital technologies.⁴²

We argue that the decision to enroll someone who may lack capacity to consent should prioritize the individual’s past wishes, personality, and identity—even if those wishes may not align with what others perceive as “best interest” for them in a generic sense.⁴³ This approach aligns with the underpinning system of organ donation, in which the person makes their personal wishes known before any need for such health-related decisions arise. Indeed, while not yet in common use globally, the organ donor type approach could be operationalized through advance research directives through which the person in early stages of ADRD gives explicit directives regarding future research participation when they still have capacity.⁴⁴ Efforts in this area should ensure directives are as fully informed and as inclusive as possible for potential future uses, accounting for emerging biomarkers, such as speech, alongside fast-moving technological leaps in digital and remote data collection.

5 | BARRIERS TO INCLUSION OF DIVERSE POPULATIONS

While remote assessment methods have the potential to offer more equitable health care across diverse populations, there are significant challenges in bringing these scientific advances to global communities. Cultural sensitivity is particularly crucial when engaging senior populations with varying levels of health literacy, who may hold differing perceptions of medical technology, shaped sometimes by cultural beliefs, which impact participation in clinical studies. For instance, in Taiwan, a rise in online scams and cybersecurity concerns has led to

skepticism toward participating in trials and the use of new technologies. Additionally, socioeconomic disparities pose a significant obstacle, with lower income groups being less likely to engage with advanced technologies. Our experience with DementiaBank in the United States suggests that online platforms, such as Zoom, and community outreach may yield more trusted engagement from diverse groups than traditional passive recruitment methods.⁹ However, technology adoption within the same country, for instance in the UK, as well as across other countries, remains variable. For example, in Colombia, people in rural areas may be less motivated to use health-care technology due to both a lack of technological understanding and poor connectivity, while in Brazil underserved populations tend to be more open to receiving care locally through public services and civil society organizations than remotely, through telehealth systems, for instance.

What is more, the appropriateness of cognitive or reading tasks can be heavily influenced by the literacy levels of participants. This variation in literacy can skew results, as individuals with higher literacy may perform differently on certain tasks compared to those with lower literacy levels, potentially leading to biased outcomes in the analysis of speech biomarkers, which is often the case in analyses conducted in the Colombian population.⁴⁵ A sustainable system of care for brain health that includes diverse populations means not only equipping local health-care workers with remote screening tools but also raising awareness around healthy aging in local communities and the benefits of early engagement with screening or monitoring programs for neurodegenerative disease.

6 | TRANSLATING RESEARCH ON SPEECH BIOMARKERS TO CLINICAL SETTINGS

Much of the current debate around speech-based biomarkers for use in clinical settings focuses on the lack of integration between research findings and clinical care, insufficient validation protocols for speech biomarkers, and the lack of normative data to inform clinical decisions. As a result, voice and speech assessments risk being perceived as experimental tools rather than reliable diagnostic aids in clinical settings. Arguably, a key barrier to clinical utility is limited understanding around the *output* of what speech-based assessments mean, especially compared to commonly used tools in primary or secondary care settings. A feasible approach to using speech-based biomarkers in clinical settings is therefore deploying clinician-friendly digital tools with automatic scoring and presentation of results in relation to normative data that are clear to interpret. Any patient-facing computer applications would need to be designed with patient and public involvement to aid adoption of this type of health-care technology among the target disease population.

Experience in Colombia has shown that patients value easy-to-understand immediate feedback,⁴⁵ however, varying and unknown acoustic conditions where recordings are made may generate a technical barrier for consistent output and interpretation of the results. We present some of the practical learnings in remote speech data collection and possible solutions in Table 1. Our experience in the UK

suggests that clinician training in interpreting output from health-care technology, such as remote speech collections tools, should be as quick and easy as possible to fit into clinical workflows. Importantly, any output from such digital tools would need to offer meaningful information to the clinician regarding risk of disease or progression trajectory on an individual level. We also see this in the United States, where health-care professionals' unfamiliarity with speech biomarker technology adds complexity, particularly where its benefits are unclear in the absence of or limited access to disease-modifying therapies.

Colombia's health-care system is confronting similar challenges, in which professionals in rural and remote regions, where many of the individuals who carry genetic mutations reside, are not familiar with the use of speech-based tools, which could greatly aid in early detection and ongoing management of dementia. Similarly in Greece, where specialized services are located mainly in the largest cities (Athens and Thessaloniki), a major barrier is the lack of awareness and training of general practitioners. Simple training programs could help integrate this technology into routine care.

7 | INCORPORATING SPEECH TECHNOLOGY INTO CLINICAL PRACTICE

In clinical practice, remote speech assessments can support traditional care, flag individuals who need further testing based on their ongoing monitoring, triage, and integration into workflows without requiring additional appointments at the clinic. Clinician training on interpretation of speech assessments and the integration of these outputs that go beyond raw scores into electronic health records should aid diagnosis and treatment planning. The US DementiaBank team is developing Grand Rounds online resources to provide education to students and clinicians on discourse-informed assessment and treatment that leverage the free educational resources available through TalkBank. Once clinicians feel comfortable to communicate speech-based assessment results on an individual basis (rather than just averages), remote tools could offer an additional source of certainty for the clinician regarding accuracy of diagnosis/prognosis and for the person with disease.

It is also critical to involve family members and provide education on speech biomarkers' importance, so family members can facilitate conversations in familiar interactions that generate spontaneous speech. This is particularly relevant in the UK, where most studies run by research institutions include care partner involvement for monitoring brain health. Evidence from research settings suggest that a dyad approach in which both the person with disease and their partner are involved may be preferable to some individuals when it comes to speech biomarker assessments.⁴⁶ Similarly in Greece, family may play an even greater role in dementia care because many people with AD rely on relatives for support. Therefore, involving caregivers in speech assessments could make the process more natural and improve data quality. Experience in the UK health system also suggests that opt-out rather than opt-in consent for speech data collection with transparent communication about interpreting output of speech data may yield better engagement than opt-in consents.

TABLE 1 Using speech assessments in remote monitoring of cognition.

Practical issues encountered	Mitigation strategy
Researchers do not commit to data sharing from the outset; that is, no data sharing statement in the participant informed consent form	Include an explanation about data sharing in participant information sheets and ask consent for data sharing (even if consenting to this point is optional).
Opt-in consent results in low recruitment numbers	Design study protocols with opt-out consent where appropriate.
Participants complete the remote study assessment multiple times, creating multiple entries for one data point	Consider disabling the data collection for each user ID after the assessment has been completed for a limited period, for example, 1 month (but have the platform available again for future follow-up visits). Send an e-mail confirmation after a remote visit has been completed to avoid the participant getting confused if they have completed the assessment already.
Participants complete the remote study assessment outside the visit window	Define data capture timeframe (e.g., 1 month after issuing invitation) or account for varying visit windows in longitudinal analysis (e.g., dynamic structural equation models).
Multiple participants may use the same e-mail address (e.g., a couple sharing an account), creating two entries from different individuals with the same associated e-mail address	Verify all entries with e-mail address and additional identifiers (e.g., initials, date of birth, phone number).
Participants struggle with basic software navigation	Co-develop simple set-up guides with patient or public involvement; provide technical support; identify digital champions within remote and rural communities.
Someone other than the participant performs the remote tests (e.g., grandchildren)	Implement technological safeguards, such as voice-identification algorithms. Co-design the data collection tools and instructions with input from the target population and culture. Educate health-care professionals about remote monitoring and how they can explain the importance of data collection to patients under their care.
Participants find instructions hard to follow	Adapt instructions to patients' backgrounds to ensure accessibility and comprehension. Be flexible with the battery you share with each patient depending on their ability/background.
Lack of acoustic consistency in the collected data	Ask participants to read a standardized paragraph to establish baseline acoustics and noise profiles. Encourage consistent recording conditions (hand out devices if possible to keep settings the same). Collect information about the device used (e.g., sampling rate).
Acoustic descriptors, particularly those related to spectral information, reflect the recording environment rather than to the patient's health state	Apply additional pre-processing such as segmentation of the speech data to ensure only context-independent acoustic descriptors for the patient are included.
Issues with content quality when participants fail to follow specific prompts or instructions	Co-design task prompts with the target population to ensure the prompt is interpreted in a way that produces the appropriate response; pilot the design prompts.
No rural/remote hospitals	Partner urban hospitals with regional clinics, using speech biomarkers as an early screening tool where other options are limited.
People prefer in-person health care	Integrate speech capture into routine doctor visits instead of relying solely on remote methods; combine digital delivery with in-person delivery of remote data collection and care.

In countries such as Brazil and Colombia, which cover large geographic areas and are characterized by uneven population distributions and regional diversity, remote assessment is also seen as an opportunity for better understanding of the epidemiological profile of ADRD. However, there may be operational difficulties in individuals completing speech-based assessments remotely, despite providing both voice and text instructions. Across the settings with which the authors of this paper have experience, we commonly observe that participants struggle with basic software navigation, and after multiple failed attempts, many simply give up in frustration.

On occasions when individuals have been unable to complete the tasks themselves, individuals have asked their grandchildren to perform the remote tests—a well-intentioned but problematic solution that compromises data integrity. This type of data collection therefore needs careful consideration around privacy and task authenticity in shared living environments. Even when limited to assisting the individual taking the test, dependence on grandchildren setting up the assessment has introduced a bias and can affect the data. In Colombia, when the focus is on early detection and monitoring of patients,⁴⁷ it can be hard for people with an ADRD diagnosis to follow specific

instructions to perform appropriate data collection remotely,⁴⁵ compounded by noisy environments, which can degrade the quality of speech recordings. Additionally, in the UK we have observed that many senior couples may share a single e-mail address between two people, which adds a layer of complication to remote data collection, particularly if platforms are designed to prevent duplicate entries with the same e-mail address to prevent the same participant accidentally completing the assessment more than once. These practical experiences underline a need for simple, user-friendly platforms.

8 | PATHWAY TOWARD VALIDATION OF SPEECH BIOMARKERS

One of the limitations of speech biomarkers in AD and their limited uptake in clinical settings is a lack of clinical validation. Development of clinically meaningful insights from speech biomarkers would require interdisciplinary collaboration between data analysts and clinical experts, robust data collection at a large scale, and alignment with gold-standard neurodegeneration metrics. Validation of speech biomarkers should involve the use of established biomarkers, such as amyloid positron emission tomography or cerebrospinal fluid analysis, as a reference standard, though there remains uncertainty around the psychometric properties even with gold standard biomarker assessments, primarily regarding specificity of the underlying etiology of disease pathology.⁴⁸

While standardization of how data are collected or harmonized across existing data sets would ensure generalizability across populations,⁴⁹ validation strategies should also be shaped by cultural contexts. In Taiwan, for example, where there is a strong cultural preference to receive health care in person, a window of engagement has been during the extended waiting times for in-person appointments when speech data could be collected. In the United States, a move toward decentralized trials has resulted in an increase in mobile-enabled tools to reach diverse populations. In Brazil, the structure of the public health-care system also favors decentralization. Validating speech biomarkers for AD in Greece and Colombia comes with challenges but also exciting opportunities that complement traditional assessment methods. Given the scarcity of non-English datasets, collecting language-specific data is crucial for accurate modelling and future results⁵⁰.

9 | EMPOWERING INDIVIDUALS TO MAKE DECISIONS ABOUT THEIR BRAIN HEALTH

In a real-world setting, difficulties in language and communication may be dismissed as part of normal aging, or “stress related” based on our experience, for example, in Taiwan. However, robust speech assessments in AD and comparison to normative data may result in high ecological validity and meaningful insights regarding the likely underlying causes of subtle changes observed in real-life speech. Additionally, knowledge of underlying causes may lead to greater engagement with

clinical services and monitoring of brain health. Despite the barriers outlined above in adopting new health technologies, we speculate that *willingness* to engage with speech-based assessments might be due to these tools’ alignment with everyday behaviors, and that this should be leveraged for positive engagement.

Evidence around disclosure of early brain disease like MCI suggests clinicians require better tools to identify, disclose, and monitor clinical syndrome and neurodegenerative disease diagnoses.⁵¹ In this regard, insights and metrics from remotely conducted speech-based monitoring could be used to support meaningful conversations about disease status and follow-up care between the clinician and the person engaging with services. We note that output from cognitive assessments using speech-based assessments should be explained to the person in the same manner as physical measurements (blood pressure, cholesterol levels, etc.), with interpretation of trends over time and of what these indicate, predict, and correlate with. It is critical that detection and monitoring methods enable a person to make informed decisions, whether that is the decision to enroll in clinical trials, consider taking newly approved treatments, or modify lifestyle to reduce risk of further decline. In cultural contexts globally, where subtle speech changes are often normalized in older age, speech biomarkers shift perceptions from accepting certain decline to empowering individuals to maintain cognitive health and encourage early intervention.

There is no consensus on whether the person completing speech-based assessments remotely should have access to their own ongoing monitoring results, or indeed, whether their care partner should have such access. Although there are concerns that cognitive health assessments may highlight declines that could negatively impact patients’ emotional well-being if appropriate clinical management is not in place, our experience in Colombia and Brazil has shown that older adults *want* to know about their health/brain state and are likely to look for technology to provide them with feedback and further recommendations to keep improving their health. Indeed, empowering individuals with knowledge of their health seems most appropriate, especially as poor performance may flag either a separate modifiable condition or allow individuals to take action to slow down progression of the neurodegenerative process. In those cases, health technology could be used to create new speech/language tasks that motivate keeping mentally active. However, for such a technology to be used worldwide, it is necessary to solve the problem of creating robust AI foundation models in different languages, especially adapted to different neurological conditions like ADRD.

10 | FUTURE DIRECTIONS

The focus for research and development programs advancing speech technologies in neurodegenerative disease should be on collecting data from diverse populations, and ensuring these data are widely accessible to speech biomarker researchers. Data sharing should include not only the development of technical means of privacy protection adequate for protection of speech data, but also ethical and legal frameworks for data collection and sharing. Addressing these issues

is crucial to ensure current barriers to biomarker validation are overcome, and subsequent translation to clinical practice is enabled. To date, no large-scale studies exist on the use of speech as a means of cognitive assessment at a public and global health level. There is a need for larger, global health research programs to be implemented to investigate the role of speech in promoting effective public health approaches to meet the WHO ADRD diagnosis targets and ultimately empowering individuals to make decisions about their brain health.

11 | CONCLUSION

Speech biomarkers have great potential as a cost-effective and non-invasive method for engaging individuals remotely for the detection and monitoring of neurodegenerative disease behavioral markers. In this paper, we draw on our experience from a range of locations across the world to highlight the diagnostic potential, limitations, and future directions for the speech biomarker field in neurodegenerative disease. While our paper does not intend to be exhaustive of these areas, we illustrate that there are shared challenges in the adoption of speech biomarkers in monitoring cognition and brain health. The main areas that need solutions include acquiring multilingual data sets, standardization in data collection and analysis, privacy (accounting for varying regional regulatory frameworks), and interdisciplinary collaboration. In particular, interdisciplinary collaboration should include not only disciplines such as machine learning, signal processing, neurology, and speech and language therapy, but also linguistics and cultural anthropology to ensure that local languages and cultures are appropriately addressed. Finally, a consensus on the clinical interpretation of speech biomarker data is also needed.

ACKNOWLEDGMENTS

We thank the reviewers of earlier versions of this paper. S Luz acknowledges support from UKRI grant number 10102226, for participation in Horizon Europe Project INT-ACT.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the [Supporting Information](#).

ORCID

Stina Saunders  <https://orcid.org/0000-0003-3323-2505>

Saturnino Luz  <https://orcid.org/0000-0001-8430-7875>

REFERENCES

- Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of speech-based digital biomarkers: review and recommendations. *Digit Biomark*. 2020;4(3):99-108. doi:10.1159/000510820
- de la Fuente García S, Ritchie CW, Luz S. Protocol for a conversation-based analysis study: PREVENT-ED investigates dialogue features that may help predict dementia onset in later life. *BMJ*. 2019;9(3):e026254. doi:10.1136/bmjopen-2018-026254
- Albert P, Haider F, Luz S. CUSCO: an unobtrusive custom secure audio-visual recording system for ambient assisted living. *Sensors*. 2024;24(5). doi:10.3390/s24051506
- Rusz J, Krack P, Tripoliti E. From prodromal stages to clinical trials: The promise of digital speech biomarkers in Parkinson's disease. *Neurosci Biobehav Rev*. 2024;167:105922. doi:10.1016/j.neubiorev.2024.105922
- De Silva U, Madanian S, Olsen S, et al. Clinical decision support using speech signal analysis: systematic scoping review of neurological disorders. *J Med Internet Res*. 2025;27:e63004. doi:10.2196/63004
- van den Berg RL, de Boer C, Zwan MD, et al. Digital remote assessment of speech acoustics in cognitively unimpaired adults: feasibility, reliability and associations with amyloid pathology. *Alzheimers Res Ther*. 2024;16(1):176. doi:10.1186/s13195-024-01543-3
- Cho S, Olm CA, Ash S, et al. Automatic classification of AD pathology in FTD phenotypes using natural speech. *Alzheimers Dement*. 2024;20(5):3416-28. doi:10.1002/alz.13748
- Angelopoulou E, Papageorgiou SG. Telemedicine in Alzheimer's disease and other dementias: where we are?. *J Alzheimers Dis*. 2025;103(1):3-18. doi:10.1177/13872877241298295
- Lanzi AM, Saylor AK, Fromm D, Liu H, MacWhinney B, Cohen ML. DementiaBank: theoretical rationale, protocol, and illustrative analyses. *Am J Speech Lang Pathol*. 2023;32(2):426-38. doi:10.1044/2022-AJSLP-22-00281
- Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. *Biol Psychiatry*. 2012;72(7):580-7. doi:10.1016/j.biopsych.2012.03.015
- König A, Linz N, Zeghari R, et al. Detecting apathy in older adults with cognitive disorders using automatic speech analysis. *J Alzheimers Dis*. 2019;69(4):1183-93.
- Saeedi S, Hetjens S, Grimm MOW, Barsties VLB. Acoustic speech analysis in Alzheimer's disease: a systematic review and meta-analysis. *J Prev Alzheimers Dis*. 2024;11(6):1789-97. doi:10.14283/jpad.2024.132
- Boahen O, Owusu-Agyei S, Fehir LG, et al. Community perception and beliefs about blood draw for clinical research in Ghana. *Trans R Soc Trop Med Hyg*. 2013;107(4):261-5. doi:10.1093/trstmh/trt012
- Sylla B, Ismaila O, Diallo G. 25 years of digital health toward universal health coverage in low- and middle-income countries: rapid systematic review. *J Med Internet Res*. 2025;27:e59042. doi:10.2196/59042
- World Health Organization. Global status report on the public health response to dementia. <https://www.who.int/publications/item/9789240033245>
- Fristed E, Skirrow C, Meszaros M, et al. Leveraging speech and artificial intelligence to screen for early Alzheimer's disease and amyloid beta positivity. *Brain Commun*. 2022;4(5):fcac231. doi:10.1093/braincomms/fcac231
- Hajjar I, Okafor M, Choi JD, et al. Development of digital voice biomarkers and associations with cognition, cerebrospinal biomarkers, and neural representation in early Alzheimer's disease. *Alzheimers Dement*. 2023;15(1):e12393.
- Pernecky R, Jessen F, Grimmer T, et al. Anti-amyloid antibody therapies in Alzheimer's disease. *Brain*. 2023;146(3):842-9. doi:10.1093/brain/awad005
- García AM, de Leon J, Tee BL, Blasi DE, ML Gorno-Tempini. Speech and language markers of neurodegeneration: a call for global equity. *Brain*. 2023;146(12):4870-9.
- Yang M, El-Attar AA, Chaspari T. Deconstructing demographic bias in speech-based machine learning models for digital health. *Front Digit Health*. 2024;6:1351637. doi:10.3389/fdgh.2024.1351637
- Ali Meerza SI, Li Z, Liu L, Zhang J, Liu J. Fair and privacy-preserving Alzheimer's disease diagnosis based on spontaneous speech anal-

- ysis via federated learning. *Annu Int Conf IEEE Eng Med Biol Soc.* 2022;2022:1362-5.
22. Dutta M, Mello TMD, Cheng Y, et al. Universal and language-specific connected speech characteristics of bilingual speakers with alzheimer's disease: insights from case studies of structurally distinct languages. *J Speech Lang Hear Res.* 2024;67(4):1143-64. doi:10.1044/2024_JSLHR-23-00254
 23. Bose A, Dutta M, Dash NS, Nandi R, Dutt A, Ahmed S. Importance of task selection for connected speech analysis in patients with Alzheimer's disease from an ethnically diverse sample. *J Alzheimers Dis.* 2022;87(4):1475-1481. doi:10.3233/JAD-220166
 24. Choudhury M. Generative AI has a language problem. *Nat Hum Behav.* 2023;7(11):1802-3. doi:10.1038/s41562-023-01716-4
 25. Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. *Nature.* 2023;616(7956):259-65. doi:10.1038/s41586-023-05881-4
 26. Kröger JL, Lutz OH-M, Raschke P. Privacy implications of voice and speech analysis—information disclosure by inference. In: Friedewald M, Önen M, Lievens E, Krenn S, Fricker S, eds. *Privacy and Identity Management. Data for Better Living: AI and Privacy. Privacy and Identity.* Springer International Publishing; 2019:242-258.
 27. Vakulenko S, Md Rijke, Cochez M, Savenkov V, Polleres A, Measuring semantic coherence of a conversation. In: *The Semantic Web – ISWC 2018: 17th International Semantic Web Conference, Monterey, CA, USA, October 8-12, 2018, Proceedings, Part I.* Springer-Verlag; 2018: 634-651.
 28. Chen M, Lu L, Wang J, et al. VoiceCloak: adversarial example enabled voice de-identification with balanced privacy and utility. *Proc ACM Interact Mob Wearable Ubiquitous Technol.* 2023;7(2):48. doi:10.1145/3596266
 29. Fagherazzi G, Fischer A, Ismael M, Despotovic V. Voice for health: the use of vocal biomarkers from research to clinical practice. *Digit Biomark.* 2021;5(1):78-88. doi:10.1159/000515346
 30. Tayebi Arasteh SR-UC, Nöth E, Maier A, Yang SH, Rusz J, Orozco-Arroyave JR. Federated learning for secure development of AI models for Parkinson's disease detection using speech from different languages. *Proc Interspeech.* 2023:5003-7. doi:10.21437/Interspeech.2023-2108
 31. Abadi M, Chu A, Goodfellow I, et al. Deep learning with differential privacy. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.* Association for Computing Machinery; 2016: 308-18. doi:10.1145/2976749.2978318
 32. Tobon C. Data privacy laws in Latin America: an overview. *Int'l L News.* 2015;44:1.
 33. Dourado DDA, Aith FMA. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. *Revista de Saúde Pública.* 2022;56(80). doi:10.11606/s1518-8787.2022056004461
 34. Dourado DDAaA FMA. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. *Revista de Saúde Pública.* 2022;56:80. doi:10.11606/s1518-8787.2022056004461
 35. Lalli MA, Cox HC, Arcila ML, et al. Origin of the PSEN1 E280A mutation causing early-onset Alzheimer's disease. *Alzheimers Dement.* 2014;10 suppl(3):S277-S83.e10.
 36. Ibanez A, Yokoyama JS, Possin KL, et al. The Multi-partner consortium to expand dementia research in Latin America (ReDLat): driving multicentric research and implementation science. *Front Neurol.* 2021;12. doi:10.3389/fneur.2021.631722
 37. Duran-Aniotz C, Sanhueza J, Grinberg LT, et al. The Latin American Brain Health Institute, a regional initiative to reduce the scale and impact of dementia. *Alzheimers Dement.* 2022;18(9):1696-8. doi:10.1002/alz.12710
 38. Toga AW, Phatak M, Pappas I, et al. The pursuit of approaches to federate data to accelerate Alzheimer's disease and related dementia research: GAAIN, DPUK, and ADDI. *Front Neuroinform.* 2023;17:1175689. doi:10.3389/fninf.2023.1175689
 39. Largent EA, Stites SD, Harkins K, Karlawish J. That would be dreadful: the ethical, legal, and social challenges of sharing your Alzheimer's disease biomarker and genetic testing results with others. *J Law Biosci.* 2021;8(1). <https://doi.org/10.1093/jlb/lsab004>
 40. Scottish Executive Social Research. *Adults with incapacity (Scotland) Act.* Scottish Government. 2008
 41. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. *ICH Harmonized Tripartite Guideline. [Accessed 12 May 2025]; Guideline for Good Clinical Practice E6(R1) 1996 Available at <https://www.ich.org/page/efficacy-guidelines><https://www.ich.org/page/efficacy-guidelines>*
 42. Hanrahan M, Wilson C, Keogh A, et al. How can patients shape digital medicine? A rapid review of patient and public involvement and engagement in the development of digital health technologies for neurological conditions. *Expert Rev Pharmacoecon Outcomes Res.* 2025;25(2):137-54. doi:10.1080/14737167.2024.2410245
 43. Yarborough M. Adults are not big children: examining surrogate consent to research using adults with dementia. *Camb Q Health Ethics.* 2002;11(2):160-8. doi:10.1017/S0963180102112096
 44. Jongsma KR, van de Vathorst S. Beyond competence: advance directives in dementia research. *Monash Bioeth Rev.* 2015;33(2-3):167-80. doi:10.1007/s40592-015-0034-y
 45. Orozco-Arroyave JR, Vásquez-Correa JC, Klumpp P, et al. *Apkinson: the smartphone application for telemonitoring Parkinson's patients through speech, gait and hands movement. Neurodegener Dis Manag.* 2020;10(3):137-57. doi:10.2217/nmt-2019-0037
 46. Ketchum FB, Chin NA, Erickson C, et al. The importance of the dyad: Participant perspectives on sharing biomarker results in Alzheimer's disease research. *Alzheimers Dement.* 2023;9:e12416.
 47. Pérez-Toro PA, Vásquez-Correa JC, Arias-Vergara T, et al. (2021). Acoustic and linguistic analyses to assess early-onset and genetic Alzheimer's disease. In: *Proceedings of ICASSP*, 8338-8342.
 48. Rabe C, Bittner T, Jethwa A, et al. Clinical performance and robustness evaluation of plasma amyloid-β(42/40) prescreening. *Alzheimers Dement.* 2023;19(4):1393-402. doi:10.1002/alz.12801
 49. Hufstедler H, Roell Y, Peña A, et al. Navigating data standards in public health: a brief report from a data-standards meeting. *J Glob Health.* 2024;14:03024. doi:10.7189/jogh.14.03024
 50. Luz S, Haider F, Fromm D, et al. An Overview of the ADReSS-M Signal Processing Grand Challenge on Multilingual Alzheimer's Dementia Recognition Through Spontaneous Speech. *IEEE Open Journal of Signal Processing.* 2024;5:738-749. doi:10.1109/OJSP.2024.3378595
 51. Saunders S, Ritchie CW, Russ TC, Muniz-Terrera G, Milne R. Assessing and disclosing test results for 'mild cognitive impairment': the perspective of old age psychiatrists in Scotland. *BMC Geriatrics.* 2022;22(1):50. doi:10.1186/s12877-021-02693-x

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Saunders S, Haider F, Lanzi AM, et al. The collection of speech data for the assessment of cognition remotely: Balancing ethical and practical challenges. *Alzheimer's Dement.* 2026;18:e70341. <https://doi.org/10.1002/dad2.70341>