Scientific Research Publishing

Health, 2026, 18(3), 221-234
https://www.scirp.org/journal/health
ISSN Online: 1949-5005
ISSN Print: 1949-4998

Macroscale Cortical Hierarchy and Spontaneous Speech for Preclinical Alzheimer's Disease: A Narrative Review Focused on Remote Screening and Longitudinal Prognosis

Zhengwei Chen, Mingtian Lu, Zhejing Ding, Tianyang Guan, Zheng Lai, Zhongliang Li, Jing Huang, Guomin Huang*, Yueling Lyu*, Yu Liu*

Cerebrovascular Diseases Department, Zhuhai Clinical Medical College of Jinan University (Zhuhai People's Hospital, The Affiliated Hospital of Beijing Institute of Technology), Zhuhai, China
Email: *guomin_huang@163.com, *cathy5527@163.com, *liuyu5680@jnu.edu.cn

How to cite this paper: Chen, Z.W., Lu, M.T., Ding, Z.J., Guan, T.Y., Lai, Z., Li, Z.L., Huang, J., Huang, G.M., Lyu, Y.L. and Liu, Y. (2026) Macroscale Cortical Hierarchy and Spontaneous Speech for Preclinical Alzheimer's Disease: A Narrative Review Focused on Remote Screening and Longitudinal Prognosis. *Health*, **18**, 221-234.
https://doi.org/10.4236/health.2026.183015

Received: February 10, 2026
Accepted: March 7, 2026
Published: March 10, 2026

Abstract

Amyloid-β (Aβ) pathology can be detected years before clinical Alzheimer's disease (AD), yet forecasting who will decline—and how rapidly—remains difficult in cognitively unimpaired (CU) and subjective cognitive decline (SCD) populations. Resting-state fMRI connectome gradients provide a low-dimensional description of macroscale cortical hierarchy, typically spanning unimodal systems to transmodal association cortex anchored in the default mode network (DMN). Alterations in gradient range, dispersion, and template similarity have been reported across the AD continuum and are increasingly described in preclinical Aβ+ cohorts. In parallel, spontaneous speech has emerged as a scalable digital phenotype: acoustic timing and pausing can be captured remotely and repeatedly, extracted automatically without manual transcription, and tracked longitudinally with minimal practice effects. This narrative review synthesizes: 1) why gradient-based hierarchy metrics are biologically plausible early functional readouts of preclinical AD, 2) the emerging evidence for atypical hierarchy in CU/SCD Aβ+ individuals, 3) the speech feature families most suitable for scalable early detection and prognosis—emphasizing timing/pause measures—and 4) practical standards for longitudinal models predicting cognitive decline slope and conversion to mild cognitive impairment (MCI). We conclude that gradients and speech—especially transcription-free timing markers—are complementary and potentially synergistic, but

translation requires rigorous confound control (motion, hearing, affect, device variability), transparent reporting, calibration, and external validation.

Keywords

Alzheimer's Disease, Default Mode Network (DMN), Digital Biomarkers, Spontaneous Speech, Resting-State fMRI

1. Introduction: The Preclinical Detection and Prognosis Gap

Biomarker advances have moved AD toward a biological definition and enabled identification of Aβ positivity well before objective impairment [1] [2]. In this review, we use "preclinical AD" within an AT(N)-style biomarker framework, where amyloid-β defines A, tau defines T, and neurodegeneration indexes downstream injury (N) [1] [2]. However, Aβ positivity alone does not uniquely determine near-term trajectories; clinical progression depends strongly on co-pathology (notably tau), resilience, and comorbidities [1]. Tau burden and topography may therefore modify the magnitude and clinical relevance of functional reorganization: hierarchy metrics and speech timing may show stronger deviations and tighter links to decline in Aβ+ individuals who are also tau-positive, whereas Aβ+ but tau-negative individuals may exhibit subtler or less prognostically informative changes. Standard neuropsychological tests can have limited sensitivity in CU cohorts, and repeated testing is affected by ceiling and practice effects, even when composites are used [3]. This motivates scalable markers that: 1) reflect distributed brain-system organization and 2) can be measured frequently and remotely.

This review focuses on two candidates with growing empirical traction:

1) Connectome gradients from rs-fMRI, which represent macroscale cortical hierarchy as continuous axes [4]-[6].

2) Spontaneous speech, especially timing and pause structure, which can be collected remotely at scale and tracked longitudinally [7]-[16].

An important theme is neurofunctional validation: linking speech features to interpretable hierarchy metrics may strengthen construct validity and improve generalizability.

To make the narrative synthesis more transparent and reproducible, we briefly summarize our scope and how evidence was selected.

Scope and selection. This is a narrative (not systematic) review. Studies were identified via targeted searches in PubMed and Google Scholar (last updated: Feb, 2026) using combinations of terms including "connectome gradient", "functional gradient", "cortical hierarchy", "default mode network", "Alzheimer*", "preclinical", "cognitively unimpaired", "subjective cognitive decline", "amyloid", "tau", "spontaneous/connected speech", "acoustic", "pause", and "remote/telephone/smartphone". We prioritized: 1) rs-fMRI gradient studies spanning the AD continuum with clearly described cohorts (biomarker-anchored when available) and

2) speech studies emphasizing timing/pause measures suitable for remote, repeated capture. We treat biomarker-anchored preclinical evidence, genetic-risk-only evidence, and symptomatic AD/MCI evidence as distinct tiers when interpreting claims.

2. Connectome Gradients and Cortical Hierarchy: What They Measure and Why They May Be Early-Sensitive

2.1. The Unimodal-to-Transmodal Axis

Gradient mapping embeds cortical regions based on similarity of their functional connectivity profiles, yielding low-dimensional axes. The principal gradient commonly spans unimodal sensory-motor cortex to transmodal association cortex, with DMN regions near the transmodal apex [4] [5] [17]-[24]. This organization is relevant to preclinical AD because Aβ burden and network vulnerability concentrate in association cortex and DMN-centered hubs [24].

2.2. Prediction-Friendly Gradient Summaries

For preclinical screening/prognosis, gradient analyses often reduce complexity into interpretable features:

1) Gradient range/distribution (e.g., "compression"),

2) Network centroid shifts (DMN/control vs sensory),

3) Dispersion in gradient space,

4) Template similarity ("fit") to canonical hierarchy [6].

Operational definitions (common practice). In practice, gradients are typically estimated by embedding a parcel-wise functional connectivity similarity matrix into a low-dimensional space and aligning individual solutions to a reference template to enable comparison [6]. "Gradient compression" generally refers to reduced dynamic range or variance of gradient coordinates (often on the principal unimodal-to-transmodal axis), consistent with weaker separation between unimodal and transmodal systems. "Dispersion" summarizes the spread of regions or networks in gradient space (e.g., within-network dispersion or between-network distances), indexing how differentiated systems are along hierarchy axes. "Template similarity/fit" quantifies how closely an individual's gradient topography matches a canonical template (e.g., spatial correlation or similarity of network ordering along the gradient), with lower fit indicating altered hierarchical organization.

These summaries can be integrated into longitudinal models without the dimensionality burden of edge-wise connectomes.

2.3. Methodological Prerequisites

Gradient features inherit rs-fMRI vulnerabilities: motion and physiology can distort connectivity structure [25] [26]; reliability of FC-derived measures is variable [27]. Multi-site work benefits from harmonization strategies and rigorous QC reporting [28] [29].

3. Preclinical Populations and Clinical Endpoints for Translation

3.1. Populations

Common targets include CU Aβ+ cohorts (CSF/PET/plasma-defined), SCD cohorts, and genetically enriched groups. Prognostic heterogeneity is expected and should be modeled rather than treated as noise.

3.2. Endpoints

For clinical relevance, studies should prioritize:

1) Cognitive decline slope (trajectory over repeated assessments),

2) Conversion to MCI (time-to-event). Diagnostic/endpoint frameworks are often based on widely used criteria and staging tools [30]-[33].

4. Evidence for Atypical Hierarchy across the AD Spectrum

To avoid over-generalizing across stages and cohort definitions, we summarize hierarchy findings in three tiers: symptomatic AD/MCI studies establish the plausibility of hierarchy reorganization; biomarker-anchored CU/SCD evidence provides the strongest support for preclinical sensitivity; and genetic-risk-only studies suggest vulnerability signals but are not pathology-specific without biomarker confirmation.

4.1. Symptomatic AD/MCI Evidence (Context and Plausibility)

Hierarchy and network-organization differences are reported in symptomatic AD and MCI, motivating early-stage work and providing face validity for hierarchy-based markers [34].

4.2. Biomarker-Anchored CU/SCD Aβ+ Evidence (Preclinical Focus)

Biomarker-anchored preclinical studies have reported gradient properties related to CSF biomarker profiles and memory performance in asymptomatic at-risk cohorts, supporting biomarker sensitivity of hierarchy metrics in the absence of overt impairment [35].

A key CU cohort study (ALFA+) combined CSF-defined Aβ status, whole-cortex gradients, and telephone-recorded spontaneous speech. CU Aβ+ participants showed widespread gradient alterations and a "compressed" gradient distribution, while a higher-order hierarchy metric related specifically to speech timing measures rather than pitch-based measures [36]. This triangulates Aβ status → hierarchy reconfiguration → scalable timing behavior, even when conventional neuropsychological separation is small.

4.3. Genetic-Risk-Only Evidence (Interpretation Limits)

Genetic-risk-only cohorts (e.g., family history or APOE enrichment without biomarker confirmation) can show gradient differences consistent with vulnerabil-

ity, but these effects cannot be attributed to AD pathology per se without Aβ/tau characterization. Such evidence is therefore best used to motivate biomarker-anchored replication and to test whether gradients capture risk-related variability that precedes, and potentially interacts with, emerging AT(N) positivity [35].

5. Spontaneous Speech for Preclinical AD: Why Timing/Pause Measures Are the Pragmatic Core

5.1. Why Speech Is a Strong Candidate for Scalable Prognosis

Speech integrates lexical access, semantic selection, planning, working memory, and executive control. Crucially, it can be collected remotely, frequently, and processed automatically, enabling within-person trajectories—an advantage for forecasting slope and conversion [7]-[16].

5.2. A "Timing-First" Feature Strategy

For preclinical deployment, the most robust and scalable features are often transcription-free:

1) speaking time, articulation rate, speech rate,

2) silent pause frequency/duration and pause distribution,

3) disfluency timing patterns (filled pauses, repairs) where reliably detected.

This strategy reduces dependency on ASR accuracy and linguistic normalization.

Minimal deployable remote task set.

To translate timing-first features into scalable remote screening and longitudinal follow-up, a minimal, deployable task set should elicit timing and pausing under complementary cognitive demands.

A practical minimal battery can include: 1) a picture description task to elicit structured discourse under planning demands and enable stable timing/pause quantification in widely used benchmarking contexts [13] [14]; 2) a brief autobiographical prompt to increase episodic retrieval and ecological validity, where pausing is mechanistically interpretable as production and retrieval load [37]; and 3) a short verbal fluency task (semantic or phonemic) to stress lexical retrieval and executive control, which is expected to manifest primarily as slowed output and increased pausing under retrieval pressure. Together, these tasks are feasible for phone/smartphone deployment and are compatible with repeated, transcription-free acoustic timing extraction in remote protocols [38].

5.3. Mechanistic Grounding: Pauses, Disfluency, and Cognitive Load (Psycholinguistic Foundations)

A long psycholinguistic tradition supports the interpretability of pause/disfluency as markers of planning, lexical retrieval difficulty, and cognitive load:

1) Pause length relates to predictability and production demands [39]-[41].

2) Reviews and models emphasize pauses as windows into formulation processes [42] [43].

3) Filled pauses ("uh/um") have communicative and processing consequences, not only "noise" [44]-[49].

4) Disfluencies can modulate listener expectations and comprehension online [46] [48].

These foundations justify why subtle preclinical changes might appear first in timing/pause structure, especially under increased planning or retrieval demands.

5.4. AD/MCI-Focused Pause Evidence: From Descriptive to Distributional Biomarkers

Pause-focused AD/MCI research has moved beyond mean pause duration toward pause distributions and multimodal pause characterization:

1) Pauses during autobiographical discourse reflect episodic memory processes in early AD [37].

2) Pause distribution analysis has been proposed as an early marker and tested in connected speech contexts [50].

3) Acoustic pause extraction and encoding strategies enable transcription-free AD detection pipelines [51] [52].

4) Benchmarking and reviews reinforce the prominence of timing/pause features across datasets and languages [9]-[13].

5.5. Remote Speech Cohorts, Feasibility, and "Real-World" Deployment Constraints

Remote speech collection (phone/smartphone/web) introduces device, codec, and environment variability, requiring explicit QC and reporting standards. The field is increasingly supported by:

1) Protocols and cohorts designed for remote or scalable speech capture [38].

2) Benchmark shared tasks/datasets that enable comparability and robustness testing (ADReSS/ADReSSo) [13] [14].

3) Open repositories and standardized discourse protocols (DementiaBank/TalkBank ecosystem), which facilitate reproducible feature extraction and cross-study comparison [53]-[59].

5.6. Public Datasets and Infrastructure: What Is "Available" vs "Deployable"

Key resources include:

1) TalkBank as an infrastructure for multimedia language data and tools [53].

2) AphasiaBank methods and resources that strongly influenced standardized discourse elicitation/analysis pipelines [57] [58].

3) DementiaBank standardized protocols and illustrative analyses for dementia/MCI discourse sampling [56].

4) ADReSS/ADReSSo balanced benchmark datasets for dementia detection and related prediction tasks [13] [14].

6. Neurofunctional Validation: Integrating Gradients and Speech

A major limitation of speech-only models is interpretability and specificity. Linking timing/pause features to interpretable hierarchy metrics can strengthen construct validity and help distinguish neurobiologically meaningful signals from associations driven by education, dialect, or device artifacts.

The CU Aβ+ gradient–speech linkage provides an initial template: hierarchy deviation (including reduced "fit" to canonical hierarchy) relates to speech timing measures more strongly than to pitch-based measures [36]. Building on this, the field would benefit from concrete, testable multimodal hypotheses.

Testable multimodal hypotheses and analysis approach. First, we hypothesize that greater hierarchy deviation—operationalized as stronger principal-gradient compression and/or lower template fit—will covary with slower articulation rate and longer/more frequent silent pauses, consistent with reduced segregation between transmodal DMN-centered cortex and control systems supporting efficient planning and retrieval [4] [24] [36] [42] [43]. Second, we hypothesize that network-level shifts involving DMN/control positioning in gradient space will relate more strongly to pause distributional features (e.g., a heavier right tail reflecting intermittent planning bottlenecks) than to pitch-based acoustics, aligning with prior evidence that hierarchy metrics preferentially track timing behavior [36] [50].

Analytically, these hypotheses can be tested using multivariable models that control key confounds (age, sex, education, hearing, affect, motion, and device/codec proxies) and evaluate incremental value beyond demographics/APOE/baseline cognition. In longitudinal settings, mixed-effects models and time-to-event models can quantify whether gradients, speech timing, and their interaction improve prediction of cognitive decline slope and MCI conversion, with calibration and transparent reporting [60]-[71].

7. Longitudinal Prognosis: Modeling Cognitive Slope and Conversion to MCI

7.1. Cognitive Slope

Recommended: linear mixed-effects (or Bayesian hierarchical) models to estimate individual trajectories and test whether baseline gradients and speech features add incremental predictive value beyond demographics/APOE/baseline cognition [25]-[29] [60]-[73]. Speech's feasibility for high-frequency sampling is a practical advantage for slope estimation.

7.2. Conversion to MCI

Recommended: survival analysis (Cox or flexible parametric models) with careful reporting of discrimination and calibration over clinically relevant horizons [73]. With repeated speech, joint models can treat speech features as time-varying predictors.

7.3. Evaluation Standards

Prediction work should report calibration and clinical utility (e.g., decision curve analysis), and follow modern reporting and bias-assessment guidance [60]-[64]. Avoiding leakage and optimistic bias is essential, especially with small datasets and extensive feature engineering [69] [70].

8. Confounds and Practical Recommendations (Speech Emphasis)

8.1. Hearing and Affect

Hearing loss is common in older adults and is associated with dementia risk; it can also directly influence speech timing via perceptual effort and social engagement pathways. Speech-biomarker studies should measure or adjust for hearing status when feasible [74]-[77]. Depression/anxiety can affect rate and pausing and should be assessed as covariates.

8.2. Device/Codec/Environment Variability (Remote Speech)

Remote pipelines should log device/codec/sample rate and include QC metrics (clipping, SNR proxies, speech/non-speech ratios). Timing-first features are often more robust than fine-grained voice quality metrics under compression and heterogeneous microphones.

8.3. Quantity Confounding and Task Effects

Many lexical and syntactic measures scale with output quantity and task structure; models should control for speech quantity and include task indicators when multiple elicitation paradigms are used.

9. Conclusion

Connectome gradients provide an interpretable representation of macroscale cortical hierarchy relevant to preclinical AD due to the vulnerability of transmodal association systems and DMN-centered hubs [4] [5] [24]. Emerging evidence suggests that CU Aβ+ individuals can show atypical hierarchy (including gradient compression), even when neuropsychological differences are subtle [35] [36]. Spontaneous speech—especially transcription-free timing and pause structure—offers scalable remote phenotyping, is mechanistically grounded in psycholinguistic models of production and cognitive load, and has a growing dataset and benchmarking support [9]-[13] [37]-[59]. The key next step is rigorous longitudinal validation showing incremental predictive utility for cognitive slope and MCI conversion beyond established clinical and biomarker predictors, with calibration, net-benefit analyses, and external validation [60]-[71].

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Jack, C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B., *et al.* (2018) NIA-AA Research Framework: Toward a Biological Definition of Alzheimer's Disease. *Alzheimer's & Dementia*, **14**, 535-562. https://doi.org/10.1016/j.jalz.2018.02.018

[2] Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., *et al.* (2011) Toward Defining the Preclinical Stages of Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease. *Alzheimer's & Dementia*, **7**, 280-292. https://doi.org/10.1016/j.jalz.2011.03.003

[3] Donohue, M.C., Sperling, R.A., Salmon, D.P., Rentz, D.M., Raman, R., Thomas, R.G., *et al.* (2014) The Preclinical Alzheimer Cognitive Composite: Measuring Amyloid-related Decline. *JAMA Neurology*, **71**, 961-970. https://doi.org/10.1001/jamaneurol.2014.803

[4] Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., *et al.* (2016) Situating the Default-Mode Network along a Principal Gradient of Macroscale Cortical Organization. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 12574-12579. https://doi.org/10.1073/pnas.1608282113

[5] Huntenburg, J.M., Bazin, P. and Margulies, D.S. (2018) Large-Scale Gradients in Human Cortical Organization. *Trends in Cognitive Sciences*, **22**, 21-31. https://doi.org/10.1016/j.tics.2017.11.002

[6] Vos de Wael, R., Benkarim, O., Paquola, C., Lariviere, S., Royer, J., Tavakol, S., *et al.* (2020) BrainSpace: A Toolbox for the Analysis of Macroscale Gradients in Neuroimaging and Connectomics Datasets. *Communications Biology*, **3**, Article No. 103. https://doi.org/10.1038/s42003-020-0794-7

[7] Bethlehem, R.A.I., Paquola, C., Seidlitz, J., Ronan, L., Bernhardt, B., Consortium, C., *et al.* (2020) Dispersion of Functional Gradients across the Adult Lifespan. *NeuroImage*, **222**, Article ID: 117299. https://doi.org/10.1016/j.neuroimage.2020.117299

[8] He, R., Al-Tamimi, J., Sánchez-Benavides, G., Montaña-Valverde, G., Domingo Gispert, J., Grau-Rivera, O., *et al.* (2024) Atypical Cortical Hierarchy in Aβ-Positive Older Adults and Its Reflection in Spontaneous Speech. *Brain Research*, **1830**, Article ID: 148806. https://doi.org/10.1016/j.brainres.2024.148806

[9] Veréb, D., Mijalkov, M., Chang, Y., Canal-Garcia, A., Gomez-Ruis, E., Maass, A., *et al.* (2023) Functional Gradients of the Medial Parietal Cortex in a Healthy Cohort with Family History of Sporadic Alzheimer's Disease. *Alzheimer's Research & Therapy*, **15**, Article No. 82. https://doi.org/10.1186/s13195-023-01228-3

[10] Hu, Q., Li, Y., Wu, Y., Lin, X. and Zhao, X. (2022) Brain Network Hierarchy Reorganization in Alzheimer's Disease: A Resting-State Functional Magnetic Resonance Imaging Study. *Human Brain Mapping*, **43**, 3498-3507. https://doi.org/10.1002/hbm.25863

[11] Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L. and Petersen, S.E. (2012) Corrigendum to "Spurious but Systematic Correlations in Functional Connectivity MRI Networks Arise from Subject Motion". *Neuroimage*, **59**, 2142-2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

[12] Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., *et al.* (2017) Benchmarking of Participant-Level Confound Regression Strategies for the Control of Motion Artifact in Studies of Functional Connectivity. *NeuroImage*, **154**, 174-187. https://doi.org/10.1016/j.neuroimage.2017.03.020

[13] Noble, S., Scheinost, D. and Constable, R.T. (2019) A Decade of Test-Retest Reliability of Functional Connectivity: A Systematic Review and Meta-Analysis. *NeuroImage*, **203**, Article ID: 116157. https://doi.org/10.1016/j.neuroimage.2019.116157

[14] Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., *et al.* (2018) fMRIPrep: A Robust Preprocessing Pipeline for Functional MRI. *Nature Methods*, **16**, 111-116. https://doi.org/10.1038/s41592-018-0235-4

[15] Yu, M., Linn, K.A., Cook, P.A., Phillips, M.L., McInnis, M., Fava, M., *et al.* (2018) Statistical Harmonization Corrects Site Effects in Functional Connectivity Measurements from Multi-Site fMRI Data. *Human Brain Mapping*, **39**, 4213-4227. https://doi.org/10.1002/hbm.24241

[16] Collins, G.S., Moons, K.G.M., Dhiman, P., Riley, R.D., Beam, A.L., Van Calster, B., *et al.* (2024) TRIPOD + AI Statement: Updated Guidance for Reporting Clinical Prediction Models That Use Regression or Machine Learning Methods. *BMJ*, **385**, e078378. https://doi.org/10.1136/bmj-2023-078378

[17] Wolff, R.F., Moons, K.G.M., Riley, R.D., Whiting, P.F., Westwood, M., Collins, G.S., *et al.* (2019) PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Annals of Internal Medicine*, **170**, 51-58. https://doi.org/10.7326/m18-1376

[18] Moons, K.G.M., Damen, J.A.A., Kaul, T., Hooft, L., Andaur Navarro, C., Dhiman, P., *et al.* (2025) PROBAST + AI: An Updated Quality, Risk of Bias, and Applicability Assessment Tool for Prediction Models Using Regression or Artificial Intelligence Methods. *BMJ*, **388**, e082505. https://doi.org/10.1136/bmj-2024-082505

[19] Vickers, A.J. and Elkin, E.B. (2006) Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Medical Decision Making*, **26**, 565-574. https://doi.org/10.1177/0272989x06295361

[20] Pencina, M.J., D' Agostino, R.B., D' Agostino, R.B. and Vasan, R.S. (2007) Evaluating the Added Predictive Ability of a New Marker: From Area under the ROC Curve to Reclassification and Beyond. *Statistics in Medicine*, **27**, 157-172. https://doi.org/10.1002/sim.2929

[21] DeLong, E.R., DeLong, D.M. and Clarke-Pearson, D.L. (1988) Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*, **44**, 837-845. https://doi.org/10.2307/2531595

[22] Heagerty, P.J., Lumley, T. and Pepe, M.S. (2000) Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics*, **56**, 337-344. https://doi.org/10.1111/j.0006-341x.2000.00337.x

[23] Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A. and Shulman, G.L. (2001) A Default Mode of Brain Function. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 676-682. https://doi.org/10.1073/pnas.98.2.676

[24] Buckner, R.L., Andrews-Hanna, J.R. and Schacter, D.L. (2008) The Brain's Default Network: Anatomy, Function, and Relevance to Disease. *Annals of the New York Academy of Sciences*, **1124**, 1-38. https://doi.org/10.1196/annals.1440.011

[25] Greicius, M.D., Srivastava, G., Reiss, A.L. and Menon, V. (2004) Default-Mode Network Activity Distinguishes Alzheimer's Disease from Healthy Aging: Evidence from Functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 4637-4642. https://doi.org/10.1073/pnas.0308627101

[26] Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., *et al.* (2006) Consistent Resting-State Networks across Healthy Subjects. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 13848-

13853. https://doi.org/10.1073/pnas.0601417103

[27] Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., *et al.* (2007) Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *The Journal of Neuroscience*, **27**, 2349-2356. https://doi.org/10.1523/jneurosci.5587-06.2007

[28] Thomas Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., *et al.* (2011) The Organization of the Human Cerebral Cortex Estimated by Intrinsic Functional Connectivity. *Journal of Neurophysiology*, **106**, 1125-1165. https://doi.org/10.1152/jn.00338.2011

[29] Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X., Holmes, A.J., *et al.* (2017) Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*, **28**, 3095-3114. https://doi.org/10.1093/cercor/bhx179

[30] Sheline, Y.I., Raichle, M.E., Snyder, A.Z., *et al.* (2010) Amyloid Plaques Disrupt Resting State Default Mode Network Connectivity in Cognitively Normal Elderly. *Biological Psychiatry*, **67**, 584-587. https://doi.org/10.1016/j.biopsych.2009.08.024

[31] Eyigoz, E., Mathur, S., Santamaria, M., Cecchi, G. and Naylor, M. (2020) Linguistic Markers Predict Onset of Alzheimer's Disease. *eClinicalMedicine*, **28**, Article ID: 100583. https://doi.org/10.1016/j.eclinm.2020.100583

[32] Verfaillie, S.C.J., Witteman, J., Slot, R.E.R., *et al.* (2019) High Amyloid Burden Is Associated with Fewer Specific Words during Spontaneous Speech in Individuals with Subjective Cognitive Decline. *Neuropsychologia*, **131**, 184-192. https://doi.org/10.1016/j.neuropsychologia.2019.05.006

[33] Petti, U., Baker, S. and Korhonen, A. (2020) A Systematic Literature Review of Automatic Alzheimer's Disease Detection from Speech and Language. *Journal of the American Medical Informatics Association*, **27**, 1784-1797. https://doi.org/10.1093/jamia/ocaa174

[34] Martínez-Nicolás, I., Llorente, T.E., Martínez-Sánchez, F. and Meilán, J.J.G. (2021) Ten Years of Research on Automatic Voice and Speech Analysis of People with Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article. *Frontiers in Psychology*, **12**, Article 620251. https://doi.org/10.3389/fpsyg.2021.620251

[35] König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., *et al.* (2015) Automatic Speech Analysis for the Assessment of Patients with Predementia and Alzheimer's Disease. *Alzheimer's & Dementia*: *Diagnosis*, *Assessment & Disease Monitoring*, **1**, 112-124. https://doi.org/10.1016/j.dadm.2014.11.012

[36] Fraser, K.C., Meltzer, J.A. and Rudzicz, F. (2015) Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *Journal of Alzheimer's Disease*, **49**, 407-422. https://doi.org/10.3233/jad-150520

[37] Luz, S., Haider, F., Fuente, S.d.l., Fromm, D. and MacWhinney, B. (2020) Alzheimer's Dementia Recognition through Spontaneous Speech: The ADReSS Challenge. *Interspeech* 2020, Shanghai, 25-29 October 2020, 2172-2176. https://doi.org/10.21437/interspeech.2020-2571

[38] Luz, S., Haider, F., Fuente, S.d.l., Fromm, D. and MacWhinney, B. (2021) Detecting Cognitive Decline Using Speech Only: The ADReSSo Challenge. *Interspeech* 2021, Brno, 30 August-3 September 2021, 3780-3784. https://doi.org/10.21437/interspeech.2021-1220

[39] Piau, A., Wild, K., Mattek, N. and Kaye, J. (2019) Current State of Digital Biomarker Technologies for Real-Life, Home-Based Monitoring of Cognitive Function for Mild Cognitive Impairment to Mild Alzheimer Disease and Implications for Clinical Care:

Systematic Review. *Journal of Medical Internet Research*, **21**, e12785. https://doi.org/10.2196/12785

[40] Dodge, H.H., Zhu, J., Mattek, N.C., Austin, D., Kornfeld, J. and Kaye, J.A. (2015) Use of High-Frequency In-Home Monitoring Data May Reduce Sample Sizes Needed in Clinical Trials. *PLOS ONE*, **10**, e0138095. https://doi.org/10.1371/journal.pone.0138095

[41] Pistono, A., Jucla, M., Barbeau, E.J., Saint-Aubert, L., Lemesle, B., Calvet, B., *et al.* (2016) Pauses during Autobiographical Discourse Reflect Episodic Memory Processes in Early Alzheimer's Disease. *Journal of Alzheimer's Disease*, **50**, 687-698. https://doi.org/10.3233/jad-150408

[42] Rochester, S.R. (1973) The Significance of Pauses in Spontaneous Speech. *Journal of Psycholinguistic Research*, **2**, 51-81. https://doi.org/10.1007/bf01067111

[43] Goldman-Eisler, F. (1958) Speech Production and the Predictability of Words in Context. *The Quarterly Journal of Experimental Psychology*, **10**, 96-106.

[44] Goldman-Eisler, F. (1961) The Predictability of Words in Context and the Length of Pauses in Speech. *Journal of Communication*, **11**, 95-99. https://doi.org/10.1111/j.1460-2466.1961.tb00334.x

[45] Goldman-Eisler, F. (1968) Psycholinguistics: Experiments in Spontaneous Speech. Academic Press.

[46] Levelt, W.J.M. (1989) Speaking: From Intention to Articulation. The MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001

[47] Clark, H.H. and Fox Tree, J.E. (2002) Using Uh and Um in Spontaneous Speaking. *Cognition*, **84**, 73-111. https://doi.org/10.1016/s0010-0277(02)00017-3

[48] Corley, M. and Stewart, O.W. (2008) Hesitation Disfluencies in Spontaneous Speech: The Meaning of *um*. *Language and Linguistics Compass*, **2**, 589-602. https://doi.org/10.1111/j.1749-818x.2008.00068.x

[49] Arnold, J.E., Tanenhaus, M.K., Altmann, R.J. and Fagnano, M. (2004) The Old and Thee, Uh, New: Disfluency and Reference Resolution. *Psychological Science*, **15**, 578-582. https://doi.org/10.1111/j.0956-7976.2004.00723.x

[50] Ferreira, F., Lau, E.F. and Bailey, K.G.D. (2004) Disfluencies, Language Comprehension, and Tree Adjoining Grammars. *Cognitive Science*, **28**, 721-749. https://doi.org/10.1207/s15516709cog2805_5

[51] Fraundorf, S.H. and Watson, D.G. (2011) The Disfluent Discourse: Effects of Filled Pauses on Recall. *Journal of Memory and Language*, **65**, 161-175. https://doi.org/10.1016/j.jml.2011.03.004

[52] Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J. and Pakaski, M. (2015) Speaking in Alzheimer's Disease, Is That an Early Sign? Importance of Changes in Language Abilities in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, **7**, Article 195. https://doi.org/10.3389/fnagi.2015.00195

[53] Pastoriza-Domínguez, P., Torre, I.G., Diéguez-Vide, F., Gómez-Ruiz, I., Geladó, S., Bello-López, J., *et al.* (2022) Speech Pause Distribution as an Early Marker for Alzheimer's Disease. *Speech Communication*, **136**, 107-117. https://doi.org/10.1016/j.specom.2021.11.009

[54] Liu, J., Fu, F., Li, L., Yu, J., Zhong, D., Zhu, S., *et al.* (2023) Efficient Pause Extraction and Encode Strategy for Alzheimer's Disease Detection Using Only Acoustic Features from Spontaneous Speech. *Brain Sciences*, **13**, Article 477. https://doi.org/10.3390/brainsci13030477

[55] Yuan, J., Cai, X., Bian, Y., Ye, Z. and Church, K. (2021) Pauses for Detection of Alz-

heimer's Disease. *Frontiers in Computer Science*, **2**, Article 624488. https://doi.org/10.3389/fcomp.2020.624488

[56] Gregory, S., Linz, N., König, A., Langel, K., Pullen, H., Luz, S., *et al.* (2022) Remote Data Collection Speech Analysis and Prediction of the Identification of Alzheimer's Disease Biomarkers in People at Risk for Alzheimer's Disease Dementia: The Speech on the Phone Assessment (Speak) Prospective Observational Study Protocol. *BMJ Open*, **12**, e052250. https://doi.org/10.1136/bmjopen-2021-052250

[57] Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., *et al.* (2005) The Montreal Cognitive Assessment, Moca: A Brief Screening Tool for Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, **53**, 695-699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

[58] Morris, J.C. (1993) The Clinical Dementia Rating (CDR): Current Version and Scoring Rules. *Neurology*, **43**, 2412-2414. https://doi.org/10.1212/wnl.43.11.2412-a

[59] Petersen, R.C. (2004) Mild Cognitive Impairment as a Diagnostic Entity. *Journal of Internal Medicine*, **256**, 183-194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

[60] Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., *et al.* (2011) The Diagnosis of Mild Cognitive Impairment Due to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease. *Alzheimer's & Dementia*, **7**, 270-279. https://doi.org/10.1016/j.jalz.2011.03.008

[61] Harrell Jr., F.E. (2015) Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. 2nd Edition, Springer. https://doi.org/10.1007/978-3-319-19425-7

[62] Van Calster, B., McLernon, D.J., van Smeden, M., Wynants, L. and Steyerberg, E.W. (2019) Calibration: The Achilles Heel of Predictive Analytics. *BMC Medicine*, **17**, Article No. 230. https://doi.org/10.1186/s12916-019-1466-7

[63] Riley, R.D., Snell, K.I., Ensor, J., Burke, D.L., Harrell Jr., F.E., Moons, K.G., *et al.* (2018) Minimum Sample Size for Developing a Multivariable Prediction Model: PART II—Binary and Time-To-Event Outcomes. *Statistics in Medicine*, **38**, 1276-1296. https://doi.org/10.1002/sim.7992

[64] Collins, G.S., Reitsma, J.B., Altman, D.G. and Moons, K.G.M. (2015) Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Annals of Internal Medicine*, **162**, 55-63. https://doi.org/10.7326/m14-0697

[65] Varoquaux, G. (2018) Cross-Validation Failure: Small Sample Sizes Lead to Large Error Bars. *NeuroImage*, **180**, 68-77. https://doi.org/10.1016/j.neuroimage.2017.06.061

[66] Varma, S. and Simon, R. (2006) Bias in Error Estimation When Using Cross-Validation for Model Selection. *BMC Bioinformatics*, **7**, Article No. 91. https://doi.org/10.1186/1471-2105-7-91

[67] Steyerberg, E.W. (2019) Clinical Prediction Models. 2nd Edition, Springer. https://doi.org/10.1007/978-3-030-16399-0

[68] MacWhinney, B. (2007) The Talkbank Project. In: Beal, J.C., Corrigan, K.P. and Moisl, H.L., Eds., *Creating and Digitizing Language Corpora*, Palgrave Macmillan, 163-180. https://doi.org/10.1057/9780230223936_7

[69] MacWhinney, B. (2000) The CHILDES Project: Tools for Analyzing Talk. 3rd Edition, Lawrence Erlbaum Associates.

[70] MacWhinney, B., Fromm, D., Rose, Y. and Bernstein Ratner, N. (2017) Fostering Human Rights through TalkBank. *International Journal of Speech-Language Pathology*,

20, 115-119. https://doi.org/10.1080/17549507.2018.1392609

[71] Lanzi, A.M., Saylor, A.K., Fromm, D., Liu, H., MacWhinney, B. and Cohen, M.L. (2023) DementiaBank: Theoretical Rationale, Protocol, and Illustrative Analyses. *American Journal of Speech-Language Pathology*, **32**, 426-438. https://doi.org/10.1044/2022_ajslp-22-00281

[72] MacWhinney, B., Fromm, D., Forbes, M. and Holland, A. (2011) AphasiaBank: Methods for Studying Discourse. *Aphasiology*, **25**, 1286-1307. https://doi.org/10.1080/02687038.2011.589893

[73] Forbes, M., Fromm, D. and MacWhinney, B. (2012) AphasiaBank: A Resource for Clinicians. *Seminars in Speech and Language*, **33**, 217-222. https://doi.org/10.1055/s-0032-1320041

[74] Fromm, D. and MacWhinney, B. (2016) AphasiaBank as BigData. *Seminars in Speech and Language*, **37**, 10-22. https://doi.org/10.1055/s-0036-1571357

[75] Deal, J.A., Betz, J., Yaffe, K., Harris, T., Purchase-Helzner, E., Satterfield, S., *et al.* (2016) Hearing Impairment and Incident Dementia and Cognitive Decline in Older Adults: The Health ABC Study. *The Journals of Gerontology Series A*: *Biological Sciences and Medical Sciences*, **72**, 703-709. https://doi.org/10.1093/gerona/glw069

[76] Powell, D.S., Oh, E.S., Reed, N.S., Lin, F.R. and Deal, J.A. (2022) Hearing Loss and Cognition: What We Know and Where We Need to Go. *Frontiers in Aging Neuroscience*, **13**, Article 769405. https://doi.org/10.3389/fnagi.2021.769405

[77] Loughrey, D.G., Kelly, M.E., Kelley, G.A., Brennan, S. and Lawlor, B.A. (2018) Association of Age-Related Hearing Loss with Cognitive Function, Cognitive Impairment, and Dementia: A Systematic Review and Meta-Analysis. *JAMA Otolaryngology–Head & Neck Surgery*, **144**, 115-126. https://doi.org/10.1001/jamaoto.2017.2513