P501

Automated speech analysis of a semantic verbal fluency task for the detection of cognitive impairment

Ricardo Sainz-Fuertes[1], Patricia Laura Maran[1,2,] Sergi Valero Ventura[3,4,] Berta Calm[3], Brian Murphy[5], Carla Zaldua[1]

[1]AcceXible Impacto, Sociedad Limitada, Bilbao, Spain; [2]Psychiatry, Mental Health and Addictions Group, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain; [3]Ace Alzheimer Center Barcelona, Barcelona, Spain; [4]Universitat Internacional de Catalunya, Barcelona, Spain; [5]Cumulus Neuroscience, Dublin, Ireland

1 Background

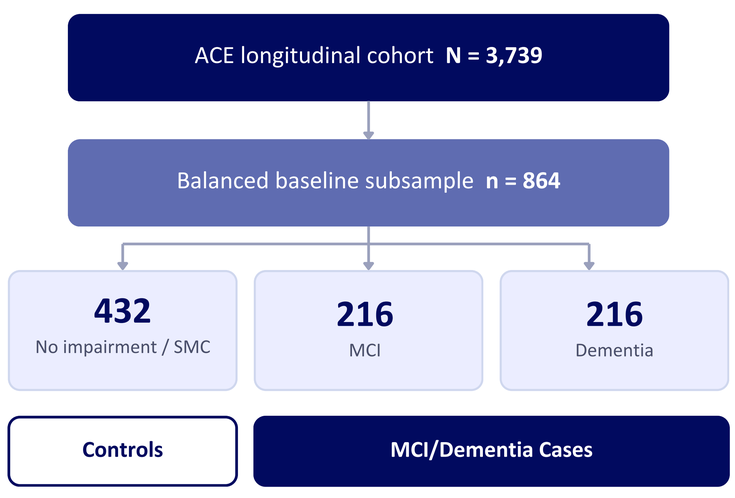
- As disease-modifying treatments arrive in the clinic, and therapy developers continue to pursue additional or adjunctive treatments, the challenge to detect early cognitive impairment grows.
- Conventional diagnostic pathways rely on extensive neuropsychological assessment, biofluid markers, and neuroimaging, which are time-consuming, costly, burdensome, and in some cases invasive, and are therefore difficult to scale to the population level.
- Speech-based digital tools offer an **accurate, quick, affordable, non-invasive alternative,** as they require no specialised equipment, take minutes to administer, and can be embedded in routine clinical workflows.
- Semantic verbal fluency (SVF) is a brief, well-established neuropsychological measure, yet its potential when analysed through automatic speech and language processing remains underexplored in large clinical samples.

Aim: This study evaluated the performance of acoustic, linguistic, and combined speech-derived models based on an SVF task (animal category) for detecting cognitive impairment in a large **Spanish-speaking** clinical sample.

2 Participants

Memory Clinic of Ace Alzheimer Center Barcelona patients, longitudinal study with **acceXible**. The present analysis used the baseline assessment only. Diagnoses established by consensus through a multidisciplinary team assessment (neurologist, neuropsychologist, social worker).

Figure 1. Study design and modelling pipeline.



Sample characteristics (n = 864)

72.2	10.4	63%
mean age, years (SD 11.0)	years of education (SD 4.9)	female

3 Speech task and models

1-minute SVF task (animal category) on a tablet via the acceXible platform, in a quiet, controlled setting under neuropsychologist supervision (for the purpose of the study only).

Acoustic 434 features; prosodic, spectral, cepstral, temporal, voice quality	→	**XGBoost**
General linguistic 12 features	→	**Random forest**
Task-specific SVF 16 features	→	**Logistic regression**

Logistic regression, random forest and XGBoost were compared within each modality; the best model per modality was retained.

Combining modalities

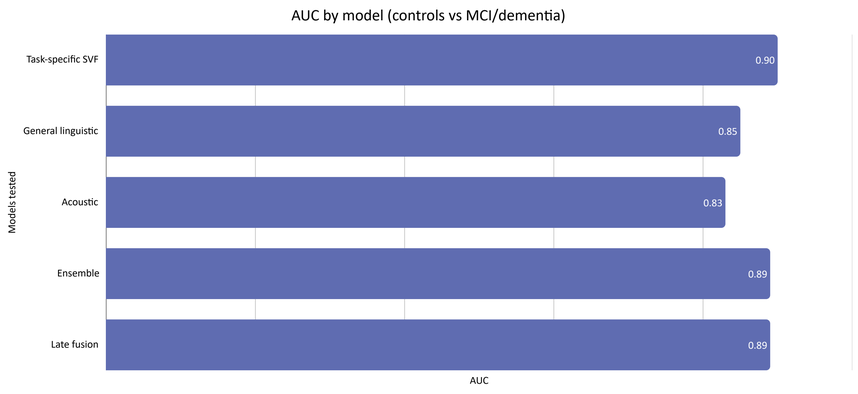
Ensemble
XGBoost on all 462 concatenated features

Late fusion
Logistic regression over the 3 modality-level predicted probabilities

4 Validation

- Five stratified outer folds; pooled out-of-fold predictions
- Fixed classification threshold of .50, no cut-off tuning
- Preprocessing, feature selection and model selection within training partitions only

5 Results



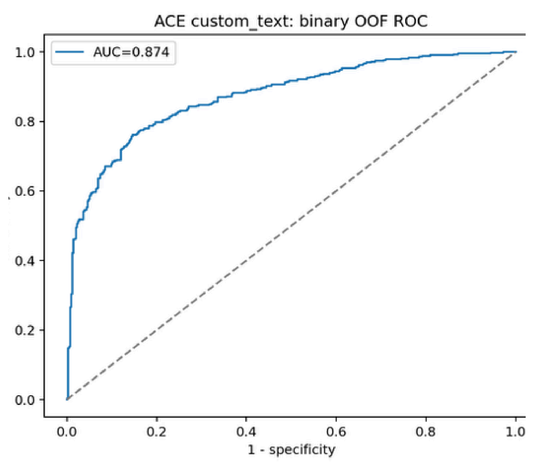
Model performance

Table 1. Performance of the classification models.

Model	AUC	Sensitivity	Specificity	Precision	F1
Task-specific SVF 16 features; LR	**0.87**	0.77	0.84	0.83	0.80
General linguistic 12 features; RF	**0.85**	0.73	0.77	0.76	0.75
Acoustic 434 features; XGBoost	**0.83**	0.76	0.76	0.76	0.76
Ensemble 462 features; XGBoost	**0.89**	0.79	0.83	0.82	0.80
Late fusion 3 probabilities; LR	**0.89**	0.79	0.83	0.82	0.81

Note. Pooled out-of-fold predictions; fixed classification threshold of .50. LR = logistic regression; RF = random forest; XGBoost = gradient-boosted trees. F1-score = harmonic mean of precision and sensitivity.

Figure 2. Receiver operating characteristic (ROC) curve of the task-specific SVF model (16 features; pooled out-of-fold predictions).



Conclusions

- A **1-minute**, tablet-administered SVF task distinguished clinically diagnosed cognitive impairment with good accuracy in a large clinic-based sample (late fusion AUC = 0.89).
- The compact task-specific feature set already implemented in the platform was the strongest single modality and captured most of the available signal; the 434 acoustic features performed least well.
- These findings support parsimonious, interpretable speech-derived measures for scalable cognitive assessment and clinical trials pre-screening.

Relevance for clinical trials

Pre-screening and cohort enrichment: a 1-minute, non-invasive, unsupervised and remotely deliverable measure for single-shot pre-screening at home or in clinic, ahead of in-person digital assessment and plasma biomarkers.

References [1] Gabirondo P, Coca A, Kaser A, et al. A novel sociodemographically-neutral speech analysis algorithm to detect cognitive impairment in a Spanish population. Alzheimer's Dement. 2023;19:e077629. [2] Gabirondo P, García-Martínez M, Pozueta-Cantudo A, Maran PL, et al. Speech biomarkers predict amyloid status in cognitively unimpaired adults. Intell Based Med. 2025;100306.

Keywords semantic verbal fluency; speech biomarkers; digital cognitive assessment; machine learning

SMC = Subjective Memory Complaints; **MCI** = Mild Cognitive Impairment

accexible.com
Scan to learn more