Machine-learning analysis of real-world multimodal data collected autonomously at home detects dementia as precisely as a traditional composite scale

Laura Rueda-Delgado, PhD (1); Alison R. Buick, PhD (2); Hugh Nolan, PhD (1); James B. Rowe, PhD, FRCP (3); Brian Murphy, PhD (1).
(1) Cumulus Neuroscience Ltd., Dublin, Ireland; (2) Cumulus Neuroscience Ltd., Belfast, United Kingdom; (3) Department of Clinical Neurosciences, University of Cambridge, United Kingdom

Cumulus NeuLogiq™ platform for use in real-world settings

Developed in collaboration with leading pharma companies and KOLs (below).

Cumulus provides full service:

- Protocol / study / SAP design
- On-site training, off-site support
- Data package
- Reporting and custom analytics

Audit ready including FDA 510(k), UKCA, HIPAA, GDPR, ISO13485.

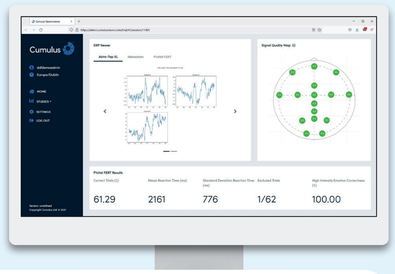
Designed for and with patients and clinicians, deployed in Phase 0-1b CNS trials.

Secure automatic upload and QC.

Real-time dashboard monitoring of decentralized and home-based data collection.

Cumulus cognitive and EEG / ERP tests are designed to be highly repeatable, with large banks of non-repeating stimuli.

- Objectively administered and automatically scored
- Results (including EEG metrics) available in minutes, enabling remote monitoring and QC
- Suitable for detecting change over time



Memory Match: visual associative memory

Symbol Swap: digit symbol substitution/coding task

Biogen · Boehringer Ingelheim · Bristol Myers Squibb · GSK · janssen · Lilly · MSD · Pfizer · Roche · Takeda

Continuous engagement with patients informs task design

Patient Advocate User Panel members are chosen for their experience with dementia, depression, schizophrenia and other neurological conditions. Each group includes a mix of patients, family members, and healthcare professionals.

Key Activities:

- Focus groups
- Usability testing
- User scenario simulation
  - In-clinic sessions
  - Remote sessions
  - Study schedules

Outcomes may influence:

- Task development
- Hardware selection
- Onboarding procedures
- Session/task list features
- Study scheduling features
- Site staff training
- Participant facing materials

FDA CLEARED · UKCA · reddot design award · iF DESIGN AWARD 2023 · GOOD DESIGN

Introduction

- Current tools for Alzheimer's disease screening and staging used in clinical research (e.g. ACE-III, ADAS-Cog) require substantial face-to-face time with trained professionals, and may be affected by subjectivity, "white coat syndrome" and other biases.
- Alzheimer's symptomology is multi-factorial and varies day-to-day. Recent advances in digital technology have enabled precise and objective measurement of cognition and function across the disease spectrum, thus potentially improving a) the sensitivity and specificity of diagnosis; and b) the ability to detect change due to an intervention, or progression of the disease.
- To enable stratification for precision treatments, more effective composite endpoints are required, with accurate quantification of functional impairment in individual domains.

Methods

Study

- We present a machine-learning stratification analysis using data from the Cumulus real-world multi-domain NeuLogiq platform, in a year-long study designed with a consortium of 10 pharma companies.
- This cross-sectional analysis utilizes 1111 sessions from repeated sampling with 101 participants (47 patients with mild Alzheimer's type dementia, and 54 healthy controls, recruited across 7 sites in the UK), during an early two-week burst period of autonomous at-home use.

Tasks

- Participants completed a range of tablet-based functional tasks with wake EEG, and separately recorded EEG during sleep.
- Tasks targeted the following functions: memory, executive function, affective processing and language.

Feature extraction

- Behaviour: number of correct responses, accuracies and reaction times from the tablet-based symbol coding task (DSST), associative memory task, n-back working memory task, visual oddball task and psychomotor task.
- Mood: accuracies and misclassification rate to anger, fear, sadness, disgust, surprise, happy faces from the facial emotion recognition task (FERT).
- Sleep: sleep quality features such as total sleep time, sleep efficiency, sleep onset latency, latency to persistent sleep, wake after sleep onset, REM latency. Provided by study partner Dreem.
- Speech: features such as global coherence, information units, syntactic complexity, errors in speech. Provided by study partner Winterlight.
- Resting state EEG: normalized theta, alpha and beta bands in frontocentral, parietal and occipital areas during eyes open and eyes close; reactivity, and exponent of log-transformed spectrum during eyes close; connectivity metrics.
- Visual oddball EEG: P300 and P100 peak amplitude and latency, event-related spectral perturbation.

ML analysis

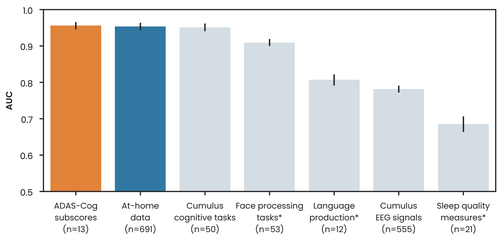
- Feature sets were used as input in separate models. Additionally, all features were used as input together in a single multi-modal model.
- Bagging of decision trees was used to compare the power of different data sources in discriminating dementia from neurotypical states. This algorithm is widely used for its performance across learning tasks (linear and non-linear), and heterogeneous datasets of differing size, levels of noise, and collinearity.
- Hyperparameter tuning was done with 10-fold cross validation.
- Model performance was assessed with 10 repetitions of the tuning.
- Feature importance from the best model was assessed by dropping variables and bootstrapping the model fit per repetition and outer fold.

Acknowledgment: The study is sponsored by Cumulus Neuroscience Ltd. and supported by Innovate UK (grant number 93826)

View this poster online

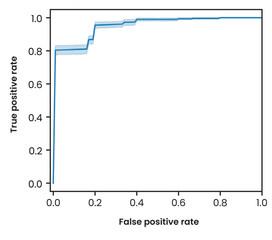
ML results

Cross-validated AUC per data type
(n: number of features)



Comparison of detection performance of ADAS-Cog versus Cumulus at-home data modalities. ADAS Cog items showed an AUC of 0.955 (SD 0.013); at-home data 0.953 (SD 0.013); Cumulus cognitive tasks 0.951 (SD 0.015); face processing tasks 0.909 (SD 0.012); language processing 0.807 (SD 0.021); EEG 0.781 (SD 0.011); Sleep quality measures 0.685 (SD 0.030). *: 3rd party assessments.

ROC curve of classifier using all at-home data



AUC plot of sensitivity/specificity performance of all at-home data modalities. 0.953 AUC, 0.88 balanced accuracy, 0.88 sensitivity and 0.86 specificity. Shaded area depicts 95% bootstrapped confidence interval.

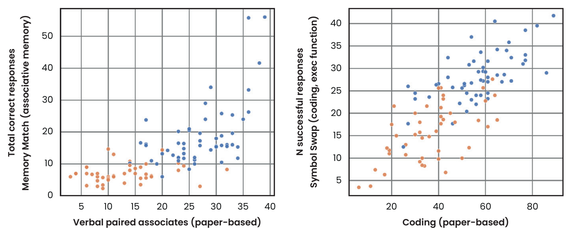
Results

Demographics and assessments

Variable	Descriptive	Dementia	Controls	Stat	Sig.	df
Sex	males:females (ratio)	30:17 (1.76)	26:28 (0.93)	1.91	n.s.	1
Education	Left formal education before age 16	9	3	7.34	n.s.	5
	Left formal education at age 16	7	6			
	Left formal education at age 17/18	10	13			
	To undergraduate degree or equivalent	11	23			
	To master degree or equivalent	6	5			
	To PhD or equivalent	4	4			
Age	mean (95%CI), n	72.8 (70.9, 74.6), n=47	70.5 (68.6, 72.4), n=54	-1.65	n.s.	99
NART	mean (95%CI), n	35.5 (32.8, 38.3), n=47	41.6 (40.4, 42.8), n=54	4.15	***	99
ADAS Cog 13	mean (95%CI), n	23.9 (21.6, 26.1), n=47	8.3 (7.2, 9.4), n=54	-12.65	***	99
Coding	mean (95%CI), n	36.9 (32.8, 41.1), n=47	56.8 (52.9, 60.7), n=54	6.79	***	99
Verbal paired associates	mean (95%CI), n	12.9 (11.2, 14.7), n=47	26.9 (25.3, 28.6), n=54	11.39	***	99
Apathy motivation index	mean (95%CI), n	1.6 (1.5, 1.7), n=47	1.2 (1.1, 1.3), n=54	-5.21	***	99
Cantril's ladder - past	mean (95%CI), n	7.0 (6.4, 7.6), n=46	8.0 (7.7, 8.3), n=54	2.8	*	98
Cantril's ladder - present	mean (95%CI), n	7.9 (7.3, 8.5), n=46	7.7 (7.2, 8.2), n=54	-0.53	n.s.	98
Cantril's ladder - future	mean (95%CI), n	5.8 (5.0, 6.6), n=46	7.5 (7.0, 8.0), n=54	3.76	**	98
Depression scale (DASS)	mean (95%CI), n	6.9 (4.9, 8.8), n=46	3.4 (2.2, 4.6), n=54	-3.07	*	98
Anxiety scale (DASS)	mean (95%CI), n	5.5 (3.8, 7.2), n=46	3.3 (2.2, 4.5), n=54	-2.14	*	98
Stress scale (DASS)	mean (95%CI), n	9.6 (7.4, 11.7), n=46	6.6 (5.1, 8.1), n=54	-2.29	*	98
Geriatric depression score	mean (95%CI), n	3.7 (2.7, 4.8), n=47	1.0 (0.6, 1.5), n=54	-4.88	***	99
Converted total score (SCI)	mean (95%CI), n	7.3 (6.6, 8.0), n=47	7.5 (6.9, 8.0), n=54	0.36	n.s.	99

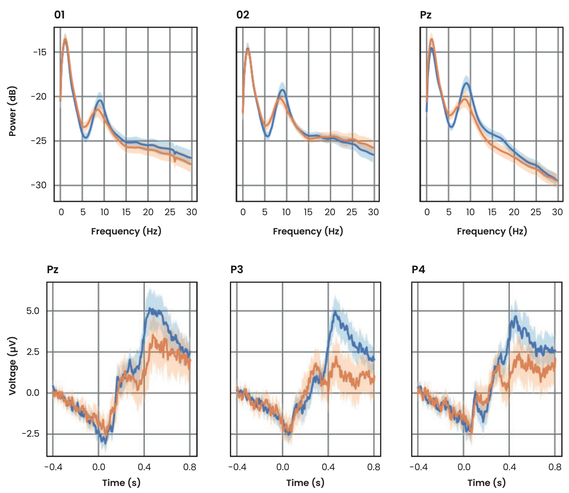
Demographics of participants included in ML analysis. Means and 95% CI of demographics and assessments per group. Non-paired t-tests were run for continuous variables, and X^2 tests were run for categorical variables. ADAS-Cog: Alzheimer's Disease Assessment Scale – Cognitive Subscale; NART: National Adult Reading Test; VPA: Verbal Paired Associates. *p<0.05; **p<0.001; ***p<0.0001; n.s: non-significant.

Correlations of Cumulus measures with benchmarks



Scatterplots of Cumulus variables against paper-based benchmarks. Left: The Cumulus' associative memory task showed a correlation with the Verbal paired associates assessment of 0.75 (p=6.2e-19). Right: The Cumulus symbol coding task showed a correlation with a paper coding assessment of 0.76 (p=5.0e-20).

Group EEG estimates



Group estimates from mixed models of the PSDs of channels O1, O2 and Pz during eyes closed. The alpha peak is reduced and slowed down, and the theta band is increased in the group with dementia when compared to the control group. 95% confidence intervals were obtained from the models' standard errors.

Group estimates from mixed models of the Target condition of channels Pz, P3 and P4. The Target condition of the visual oddball task elicited a P300 over channel Pz, with a lower amplitude in the group with dementia compared to the control group. 95% confidence intervals were obtained from the models' standard errors.

Feature	Domain	Median difference
Associative memory task: final level number	Cognitive tasks	0.00274
Associative memory task: total correct responses	Cognitive tasks	0.00234
Associative memory task: total unique patterns seen	Cognitive tasks	0.00191
Associative memory task: first four patterns % correct	Cognitive tasks	0.00177
DSST: number of successful responses	Cognitive tasks	0.00131
Oddball: P100 peak amplitude, Bonus condition, channel FCC4	EEG	0.00125
RS: Beta synchronisation likelihood within occipital area during eyes closed	EEG	0.00108
Oddball: ERSP 0.10-0.50s post-stimulus, channel P3	EEG	0.00103
RS: PSD slope during eyes closed, channel O2	EEG	0.00099
N-back: average reaction time during successful responses in match condition	Cognitive tasks	0.00095
RS: Alpha peak frequency during eyes open, channel AF8	EEG	0.00089
Sleep: respiratory rate during slow wave sleep	Sleep	0.00089
N-back: reaction time std during successful responses	Cognitive tasks	0.00086
N-back: average reaction time during successful responses to nonmatch condition	Cognitive tasks	0.00077
N-back: reaction time std during failed responses	Cognitive tasks	0.00071
Oddball: P100 peak amplitude, Target condition, channel Fz	EEG	0.00056
Oddball: P300 peak amplitude, Non Target condition, channel FCC3	EEG	0.00056
RS: Phase locking value between frontocentral and left temporal areas during eyes closed	EEG	0.00030

Top features contributing to the classifier's AUC including all at-home data. Median drop in AUC across repetitions when removing the selected feature. RS: resting state; PSD: power spectral density; ERSP: event-related spectral perturbation.

Conclusion

- Patients using digital technology autonomously in the home can yield data that matches the discriminative power of a traditional composite scale. This can enable objective precision measurement of disease severity at scale.
- Cumulus data aligns with established benchmarks, and with the literature on functional impairment in AD, constituting positive evidence for construct validity and technical feasibility.
- Usability and feasibility findings can be found in the associated poster *"A longitudinal real-world study in patients with Alzheimer's Disease dementia using frequent multi-domain digital measurements at-home performed on the Cumulus NeuLogiqTM Platform: usability and feasibility findings"* **(Diggin et al., AAIC 2024)**.