Comparison of a speech-based digital biomarker with the MoCA in a naturalistic cohort of seniors

William Simpson, PhD [1,2], Liam D Kaufman, MSc [2], Aparna Balagopalan, MSc [2,3], Jekaterina Novikova, PhD [2]

(1) Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada, (2) Winterlight Labs, Toronto, ON, Canada, (3) Department of Computer Science, University of Toronto, Toronto, ON, Canada

Background

Computational analysis of speech can be used to measure cognitive impairment and detect the presence of psychiatric and neurological disorders (1). For example, changes in speech rate, sentence length, word frequency, pronoun usage, repetitions, word finding difficulties, and number of ideas expressed are highly characteristic of Alzheimer's disease (AD) (2). The aim of this study was to compare markers extracted from computational analysis of speech against MoCA scores in a naturalistic cohort of seniors.

Methods

- 118 seniors, aged 55-95, were recruited from the community and independent living facilities in Canada and the US
- Participants completed a tablet-based assessment which included 2 picture description tasks and a MoCA was administered by a trained psychometrist.
- Using natural language processing techniques, audio samples and transcripts were analyzed to generate 9 aggregate measures pertaining to discourse, global and local coherence, information units, lexical complexity, sentiment, syntax, utterance cohesion and word finding difficulty
- Aggregate measures were expressed as z-scores against healthy norms for each metric.
- Correlations and linear regression were used to determine the relationships between aggregate measures and cognitive status

Results

Measure	Mean	SD	Range
MoCA	23.70	4.41	10-30
Discourse Mapping	-0.13	0.79	-7.4 - 1.2
Global Coherence	0.16	0.59	-1.63 - 1.31
Information Units	-0.13	0.46	-1.59 - 1.26
Lexical Complexity and Richness	-0.02	0.20	-0.64 - 0.55
Local Coherence	-0.29	0.91	-4.88 - 1.42
Sentiment	-0.42	0.84	-5.51 - 2.86
Syntactic Complexity	-0.08	0.42	-0.94 - 1.40
Utterance Cohesion	-0.37	1.23	-4.70 - 0.74
Word Finding Difficulty	0.28	0.53	-1.39 - 2.04

Figure 1: Correlations between MoCA score and aggregate measures

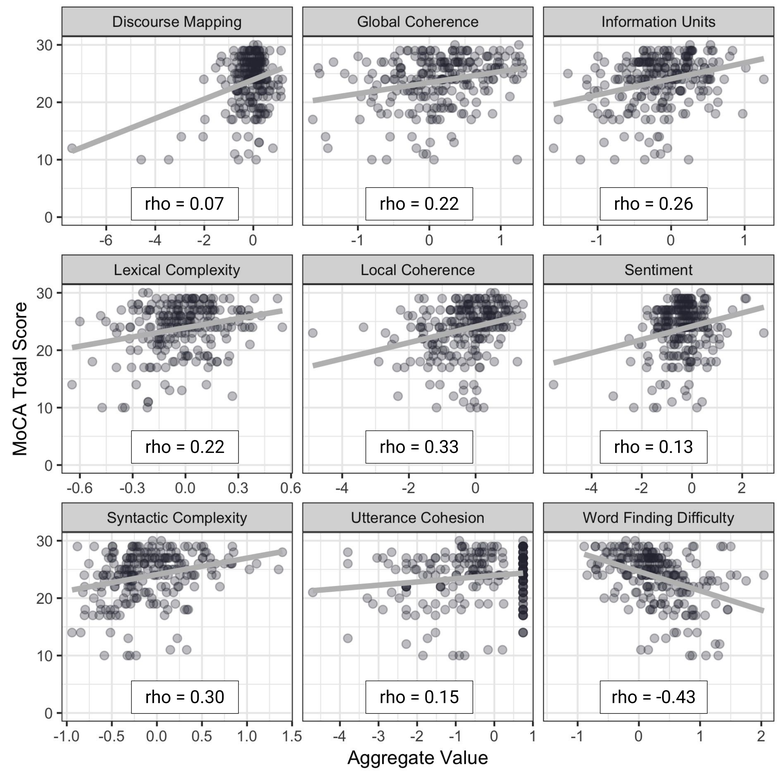
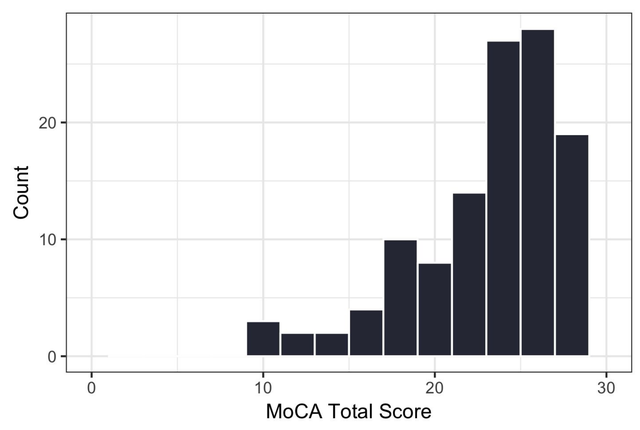


Figure 2: Distribution of MoCA Scores (n=118)



Results Continued

Model	Term	t	p
MoCA ~ all metrics Adj-R^2 = 0.307	**Discourse Mapping**	2.02	0.044
	Information Units	2.34	0.019
	Word Finding Difficulty	-2.90	0.004
	Syntactic Complexity	0.14	0.890
	Lexical Complexity and Richness	4.21	<0.001
	Global Coherence	-0.31	0.750
	Local Coherence	2.98	0.003
	Utterance Cohesion	2.03	0.043
	Sentiment	0.29	0.772

(Bold) p<0.05

Model	Term	t	p
MoCA ~ sig. metrics Adj-R^2 = 0.316	Information Units	2.58	0.010
	Local Coherence	4.02	<0.001
	Lexical Complexity and Richness	4.26	<0.001
	Word Finding Difficulty	-3.03	0.002
	Discourse Mapping	2.26	0.024
	Utterance Cohesion	2.11	0.036

(Bold) p<0.001

Conclusions

A linear regression model including each aggregate measure explained 30.7% of the variance and identified 6/9 metrics as significant predictors of MoCA score. A simplified model using only these 6 significant aggregates was similarly accurate. These results support previous work and the further development of speech-based biomarkers of cognition for use as screening and tracking tools in this population.

References

(1) Slegers, A., Filiou, R.-P., Montembeault, M. & Brambati, S. M. Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *J. Alzheimers. Dis.* 65, 519–542 (2018).

(2) Fraser, K. C., Meltzer, J. A. & Rudzicz, F. Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *J. Alzheimers. Dis.* 49, 407–422 (2016).

Presented at Alzheimer's Association International Conference (AAIC), Los Angeles, CA, USA. July 14-18, 2019