Computational linguistics analysis of cognition-related problems reported by Parkinson's Disease patients

Abhishek Hosamath[1], Lakshmi Arbatti[1], Hardik Kothare[1], Sejal Desai[1], Ira Shoulson[1] & Vikram Ramanarayanan[1,2]
[1]Modality.AI, Inc., [2]University of California, San Francisco

RATIONALE

- Prominent non-motor characteristics of Parkinson's Disease (PD) include decreased language and lexical capabilities, word-finding issues, and impaired memory.
- Free-text responses about the most bothersome PD-related problems [1] are a feature-rich source for extracting lexico-syntactic metrics [2].
- Analysis of metrics from participants reporting cognitive symptoms[3] could potentially uncover objective biomarkers of PD.

SPECIFIC AIMS

- Investigate differences in objective lexico-syntactic features of people with PD (PwPD) who did and did not report cognition issues at baseline and longitudinally.

METHODS

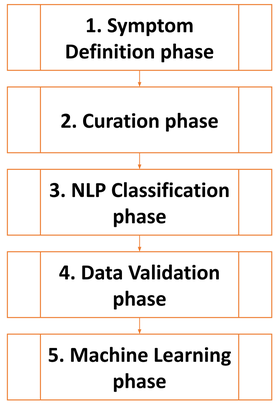
- Parkinson's Disease Patient Reports of Problems (PD-PROP) [1,3], allows PwPD to respond in their own words about their experience of up to five most bothersome PD problems to questions:

1. What is the Nth most bothersome problem due to your PD? (N=1-5)
2. In what way does this problem bother you by affecting your everyday functioning or ability to accomplish what needs to be done?

- PD-PROP data was obtained from Fox Insight (FI), an online, observational study sponsored by the Michael J. Fox Foundation (MJFF)

MJFF 2022 dataset (February 2022)	
Number of verbatim responses	380,000+
Number of participants	30,571
Age (baseline)	66±9.8
Years since diagnosis (baseline)	5±5.9
Number of PROP visits	Up to 16
Timeline of visits	4.8 years

Table 1. MJFF Cohort Characteristics



14 Motor and Non-Motor domains & 65 symptoms of PD by 9 expert curators including clinicians and PwPD[3]

Figure 1. Human-in-the-loop Curation and Classification Methodology [3]

Symptoms in Cognition domain [3]	i. Memory; ii. Language/Word Finding; iii. Concentration/Attention; iv. Cognitive Slowing; v. Executive Abilities/Working Memory; vi. Mental Alertness/Awareness; vii. Visuospatial Abilities; viii. Cognitive Impairment NOS
Lexico-syntactic features analyzed [2] (computed using spaCy)	i. word count; ii. idea density; iii. noun rate; iv. pronoun rate; v. verb rate; vi. noun-pronoun ratio; vii. noun-verb ratio; viii. closed-class-word ratio; ix. percentage content words; x. positive and negative sentiment

Baseline analysis

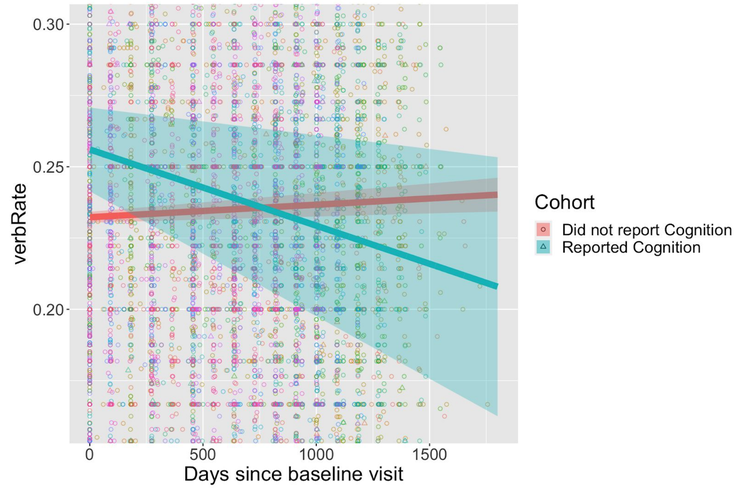
- Cohen's d was used to compare lexico-syntactic features between participants who reported Cognition as MBS (N=2611) and those who did not at baseline (N=22421)

- **Mild to moderate effects** were observed for **word count** (**0.49**) and **noun-pronoun ratio (0.34).**

Longitudinal analysis

- Growth curve models (GCM) were plotted in R for a cohort with ten visits spanning 817 – 1756.
- N=23 participants reported Cognition symptoms at all visits and N=1022 participants did not report Cognition symptoms at any of their visits.

- A statistically significant declining Verb rate (p=0.0363) was observed, consistent with multiple PD studies.



verbRate (p=0.0363; equations:
Cognition reporters verbRate = 0.00003553*days_since_baseline_visit+0.2561;
Non-Cognition reporters verbRate = 0.00000439*days_since_baseline_visit+0.2323)

Figure 2. GCM for verbRate

RESULTS AND DISCUSSION

Potentially useful PD biomarkers of Cognition

- Word count, and Noun-pronoun ratio at baseline
- Verb rate longitudinally

Limitations

- Unbalanced case-control ratio for the longitudinal analysis
- Controls could potentially report Cognition problems outside the time frame considered
- Keyboard-entry data may not be a true representation of PwPD's spontaneous speech.

REFERENCES

[1] L. Smolensky *et al.*, "Fox Insight collects online, longitudinal patient-reported outcomes and genetic data on Parkinson's disease," *Scientific Data*, vol. 7, no. 1, Art. no. 1, Feb. 2020, doi: 10.1038/s41597-020-0401-2.
[2] H. Kothare, A. Exner, S. Snyder, J. Huber, and V. Ramanarayanan, "Lexico-semantic differences between people with PD and healthy controls observed in a story retell task," *Google Docs*. https://drive.google.com/file/u/0/d/1hfSk-FiSP2xKJImcWf1_RgVqKbwsNpBs(accessed Sep. 14, 2023).
[3] C. Marras *et al.*, "What Patients Say: Large-Scale Analyses of Replies to the Parkinson's Disease Patient Report of Problems (PD-PROP)," *J Parkinsons Dis*, vol. 13, no. 5, pp. 757–767, 2023, doi: 10.3233/JPD-225083.

Acknowledgment: Research supported by Michael J Fox Foundation For Parkinson's Research (MJFF), New York.