Title: Verbatim Reports of Problems from >10,000 Parkinson Disease Patients: Informing Natural History and Clinical Outcome Assessments

Authors and Affiliations:

Shoulson, I[1]; Arbatti, L[1]; Marras C[2]; Standaert, D[3]; Tanner, CM[4]; Smolensky, L[5]; Kopil, C[5]; Hamilton, J[5]; Flagg, E[6]; Christopher CA[1]; Nguyen, A[7]

[1]Grey Matter Technologies, Sarasota, FL, USA, [2]University of Toronto, Ontario, Canada; [3]University of Alabama at Birmingham, AL, USA; [4]University of California San Francisco, CA, USA; [5]Michael J Fox Foundation for Parkinson's Disease Research, NY, NY, USA; [6] University of Rochester, NY, USA; [7] University of San Francisco, CA, USA

What is the Methodological Question Being Addressed?

What are the feasibility and utility for capturing verbatim patient reports of their problems and their functional consequences? How can clinically-curated natural language processing (NLP) and machine learning (ML) output be applied to develop: (1) a patient-reported natural history of illness (e.g., Parkinson disease), and (2) fit-for-purpose clinical outcome assessments (COAs) applied to therapeutic development of medical products?

Abstract:

Introduction: There has been no systematic research to characterize the course of neurological disease as reported directly by patients. Using natural language processing (NLP), expert clinical curation, and machine learning (ML), we analyzed the verbatim Patient Reports of Problems (PROP) from Parkinson disease (PD) patients who volunteered to answer in their own words: (1) What bothers you the most about your PD? and (2) In what way does this problem affect your daily functioning? Participants were also asked to reply to what bothers them the 2nd most, up to 5 bothersome problems.

Methods: From February-August 2018, >10,000 consenting patients worldwide entered their verbatim PROP replies, age, and date of PD diagnosis by computer keyboard (about 40% used mobile/tablet devices)

on the Michael J Fox Foundation (MJFF) FoxInsight.org research platform, having an ongoing accrual rate of about 1,000 new patients/month. Using NLP-derived and clinically-curated terms, motor and non-motor symptoms were classified with respect to age and 0-10 years since diagnosis.

Results: Motor and non-motor symptoms constituted 60% and 40% respectively of patient-reported problems: Tremor was reported at all disease durations; Rigidity and Bradykinesia were more frequent with increased duration of PD; Postural Instability (imbalance, unsteadiness) was reported early within 0-3 years of diagnosis, and was more common with age of the patient and longer duration of PD. Sleep, Pain, Mood and Cognition problems were more frequent with increased duration of PD. Fatigue symptoms were similar across duration of PD and age. Constipation was more frequent with increased age and duration of illness. These relationships will be illustrated by 3-D videos.

Conclusions: The PD-PROP represents the largest compilation of patient-reported verbatim problems for any neurological disorder. The verbatim patient reports and application of NLP/ML techniques show utility in capturing bothersome problems and provide a cross-sectional natural history of PD from the perspective of patients -- to be further informed as patients are invited every six months to update their PROP replies and longitudinal data accrue. De-identified PROP data are being made available to scientists for analysis by the MJFF.

The PD-PROP can be further customized to develop fit-for-purpose clinical outcome assessments (COAs) applicable to clinical trials. Clinically-curated and ML-informed NLP output from early-phase clinical development can be applied to characterize the PROP profile of research participants and treatment responders. These data will facilitate subject enrichment in late-phase trials and provide unfiltered and analyzable patient reports obtained at baseline and on experimental treatment or comparators (placebo).

Disclosures: This work was supported by the Michael J Fox Foundation for Parkinson's Disease Research (MJFF), NY, NY, USA, through a Pritzker Prize to I Shoulson and a Computational Science Research grant (I Shoulson, PI). All the co-authors are compensated by research funds from this MJFF funding or as scientists at the MJFF.