

Applying Natural Language Processing (NLP) to Verbatim Patient-Reported Outcomes

¹Jennifer L. Purks, ¹Michael Harris, ¹Karen E. Anderson, ¹Ira Shoulson

¹Georgetown University

Funding Support: Prana Biotechnology Ltd (Melbourne, Australia)

Financial Disclosures: Ira Shoulson, MD serves as non-executive Director of Prana Biotechnology Ltd. Karen Anderson, MD is a paid consultant to Prana Biotechnology Ltd.

Objective: To apply Natural Language Processing (NLP) to analyze verbatim patient-reported outcomes (PROs) among Huntington Disease (HD) research participants in the Phase 2 REACH2HD randomized-controlled trial examining safety and cognitive benefits of PBT2, an experimental modulator of metal ionophores.

Background: FDA and other regulatory agencies value PROs, but quantification and analytic approaches are lacking. The Huntington Disease Patient-Reported Outcome of Problems (HD-PROP) was developed to capture verbatim subject reports of patients' most bothersome problems and functional consequences. The HD-PROP was administered at baseline (BL), 12 weeks (W12) and 26 weeks (W26) of the REACH2HD trial to the 109 research participants who were randomly assigned 1:1:1 to PBT2 250 mg/day, PBT2 100 mg/day, or placebo.

Methods: Using NLP to extract relationships and meaning from large text-based resources, word clouds (images composed of words from a text where the size of the word indicates frequency or importance) were derived from HD-PROP data in the REACH2HD trial at BL and W26.

Results: The Problem Word Cloud showed that "BALANCE" (69 counts), "MEMORY" (66 counts) and "MOVEMENTS" (48 counts) were the most bothersome reported problems. There was a decrease from 8 "MEMORY" counts at BL to 4 "MEMORY" counts at W26 in the PBT2 250mg group. However the Functional Consequence Word Cloud did not show a clear or informative pattern.

Conclusion: The NLP word clouds for verbatim-reported problems and consequences help to quantify and visually depict patterns in the HD-PROP dataset of the REACH2HD trial. Verbatim problems are more uniformly informed by NLP than verbatim consequences, perhaps related to the complexity of the questions and replies. This may reflect lack of insight or impaired cause and effect reasoning between problem and its functional consequence in this patient population. Some PBT2 250mg treatment effect on MEMORY is suggested by these results. More advanced NLP may provide more informative analysis of PROs.