

## **Poster**

### **312. ALS**

**Location:** Halls B-H

**Time:** Monday, November 14, 2016, 8:00 AM - 12:00 PM

**Program#/Poster#:** 312.03/M15

**Topic:** C.05. Neuromuscular Diseases

**Support:** Barrow Neurological Foundation

**Title:** ALS and artificial intelligence: IBM watson suggests additional RNA binding proteins linked to ALS

**Authors:** \***R. P. BOWSER**<sup>1</sup>, E. ARGENTINIS<sup>2</sup>, R. SATTLER<sup>1</sup>, M. COLLINS<sup>1</sup>, A. BOEHRINGER<sup>1</sup>, I. LORENZINI<sup>1</sup>, P. FERRANTE<sup>1</sup>, A. LACOSTE<sup>2</sup>, S. SPANGLER<sup>3</sup>, N. BAKKAR<sup>1</sup>; <sup>1</sup>Barrow Neurolog. Inst., Phoenix, AZ; <sup>2</sup>IBM Watson Hlth., New York, NY; <sup>3</sup>IBM Res. - Almaden, San Jose, CA

**Abstract:** A number of RNA binding proteins (RNPs) are linked to amyotrophic lateral sclerosis (ALS), with known mutations in 11 RNPs that are causative in familial ALS and another 5 RNPs associated with ALS pathology but with no known genetic mutations. There are over 1,450 RNPs in the human genome, and therefore other RNPs may also be linked to ALS. We sought to identify additional RNPs associated with ALS using methodologies that analyze prior published

information to suggest new RNPs with a connection to ALS. The cognitive capabilities of IBM Watson enable it to extract domain specific text features from published literature to identify new connections between entities of interest, such as genes, proteins, drugs, and diseases. This approach has been successfully applied to gain new insights into oncology, but has not been applied to the neurosciences. We used IBM Watson to identify additional RNPs linked to ALS. IBM Watson analyzed published abstracts to learn the text patterns of a set of known RNPs related to ALS, and then applied that learning to a candidate set of proteins and ranked their similarity to the known RNP “training set”. To test IBM Watson’s predictive performance, we first restricted its knowledge to information prior to 2012 (known mutations in 8 RNPs linked to ALS). IBM Watson then rank ordered all other 1,445 RNPs with a probability to be linked to ALS. Since that time, mutations in three RNPs (Matrin 3, GLE1, and ARHGEF28) have been linked to familial ALS. Matrin 3 was the top candidate in this retrospective analysis, with both ARHGEF28 and GLE1 within the top 10% of all RNPs. Having shown that such analysis can successfully predict ALS- associated genes, we then applied a training set consisting of all known RNPs with mutations causative of ALS to predict other RNPs linked to ALS. Of the top 50-ranked genes, 5 have already been associated with ALS, even though no disease- causing mutations are known (RBM45, MTHFSD, SMN2, EWSR1 and hnRNPA3). Also included within the top 10 predicted genes were hnRNPU, hnRNPH2, SRSF2, SYNCRIP and CAPRIN1. To validate Watson’s

predictions, we examined the subcellular distribution of these RNPs in post-mortem tissue samples from ALS and control subjects. We identified reduced levels of SYNCRIP in ALS patients, but no changes in hnRNPH2. SYNCRIP functions in multiple steps of mRNA maturation and transport, and interacts with TDP-43, FUS, and SMN. CAPRIN1 binds G3BP and TDP-43 and accumulates in stress granules. SRSF2 is localized to antisense RNA foci in C9 patients. Overall, our approach using IBM Watson to mine scientific literature to find new ALS-linked RNPs is promising and may aid basic research efforts to understand mechanisms of disease.

**Disclosures: R.P. Bowser:** E. Ownership Interest (stock, stock options, royalty, receipt of intellectual property rights/patent holder, excluding diversified mutual funds); Iron Horse Diagnostics, Inc.. F. Consulting Fees (e.g., advisory boards); Above & Beyond, LLC.. **E. Argentinis:** None. **R. Sattler:** None. **M. Collins:** None. **A. Boehringer:** None. **I. Lorenzini:** None. **P. Ferrante:** None. **A. Lacoste:** None. **S. Spangler:** None. **N. Bakkar:** None.