Identifying Neuropsychiatric Illnesses from Voice Samples

UNIVERSITY of WASHINGTON

John Kim, Enoch Chung, Yasmin Byott, Aparna Hariprasad, Kendra Leete, Hirali Kadakia, Nikolai Liang, Christian Buensuceso, Nadav Brumer, Goda Lajauskaite, Lana Shamdeen, Kiran Pradhan, Yufei Li, Paritosh Bapat, Reza Hosseini Ghomi

PROJECT ECHO

Background

- Voice recordings have previously been able to indicate neurological/psychiatric diseases
  - Parkinson's Disease (PD) voice features include fundamental frequency, jitter, harshness, and tremor of voice (1)
  - Depression studies have analyzed fundamental frequency, pitch, intensity, and timing. Related features include jitter and shimmer, energy distribution among formants, and cepstral features (2)
  - Borderline Personality Disorder (BPD) studies have analyzed speech for pauses, word choice, and syntax (3)

Prevalence:

- PD prevalence is 0.1-0.2% of population (4)
  - Diagnosed by history & examination, and dopamine agonists (5)
- Depression prevalence is 6.7% of population (6)
  - Diagnosed by clinical interpretation of symptoms and family history (7)
- BPD prevalence is 1.4% of population (8)
  - Diagnosed by family history and interview in clinic by mental health professional (9)
- Novelty of research study:
  - Taking free voice stream samples and analyzing extracted voice features to build screening and management tool

Research Questions:

- **Can we predict if someone has PD, BPD, Depression, or no-disease based off biomarker features?**
- Proposed Hypothesis:
  - Each of the three computational models can indicate presence of specific neuropsychiatric illnesses using voice features

Methods

- Voice samples gathered alongside surveys for data comparison
- Surveys generated in REDCap utilizing standardized methods of measure for each disease
  - Modified PDQ8 (non-motor symptoms) and SPDDs (motor symptoms) for PD
  - PHQ-9 for depression
  - ZAN-BPD for BPD
- 30 seconds of free speech samples gathered with laptop or smartphone microphone using the REDCap mobile app from patients in quiet clinical environment at UW Medicine and associated clinics
- Voice samples analyzed for specific biomarkers using:
  - Random Forest (RF), Support Vector Machine (SVM), and k-nearest neighbor (kNN)
  - Measured for vocal accuracy of disease
- Each model used same voice features, but differing computational techniques for prediction
- We used the following voice features: 13 mfcc (mel frequency cepstral coefficients), jitter, shimmer, and fundamental frequency

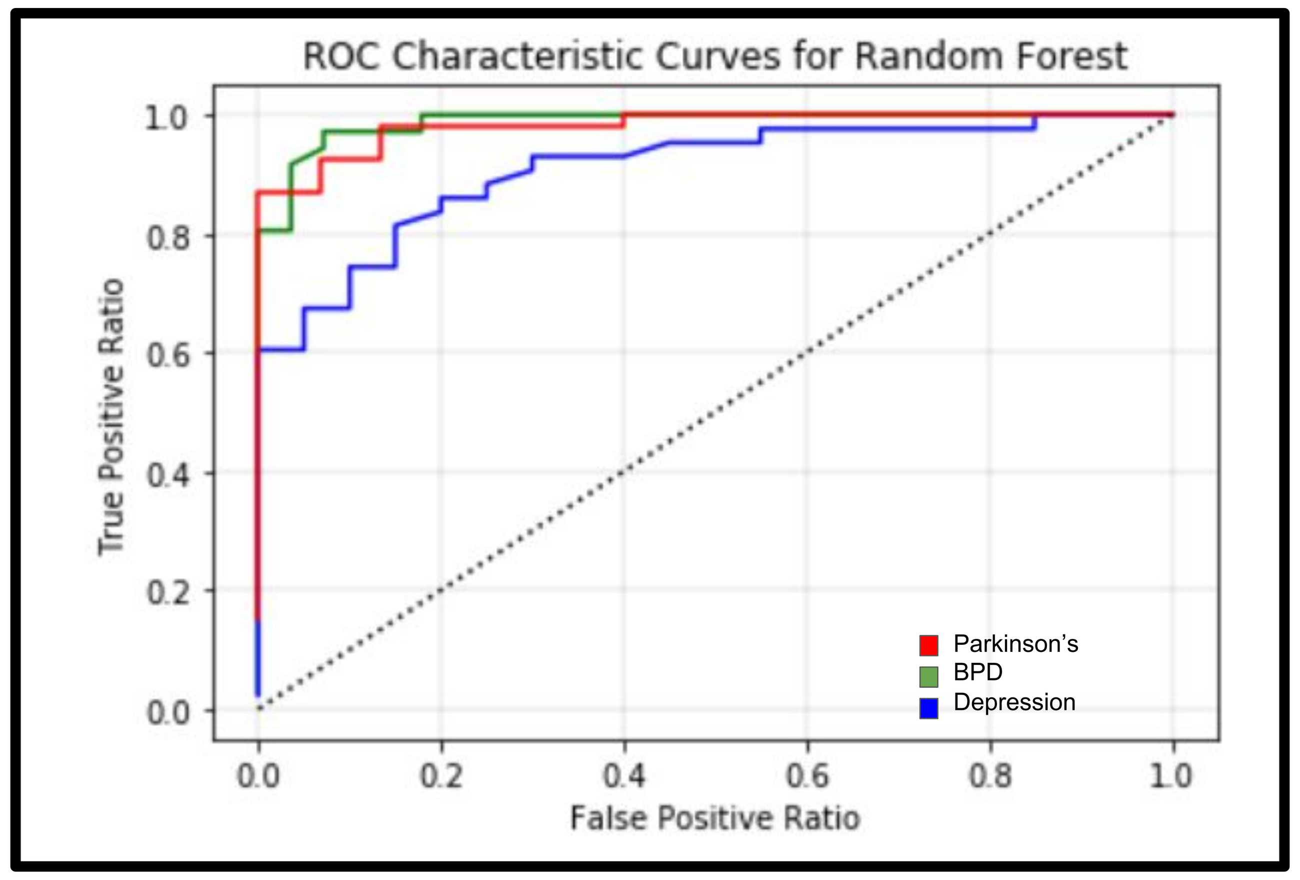

Figure 1: **Performance of Random Forest (RF) model on disease detection through voice.** Diseases are indicated by color. Dotted line indicates 50% line ratio. Curves higher on the y-axis represent greater positive accuracy (i.e. sensitivity and specificity for disease presence).

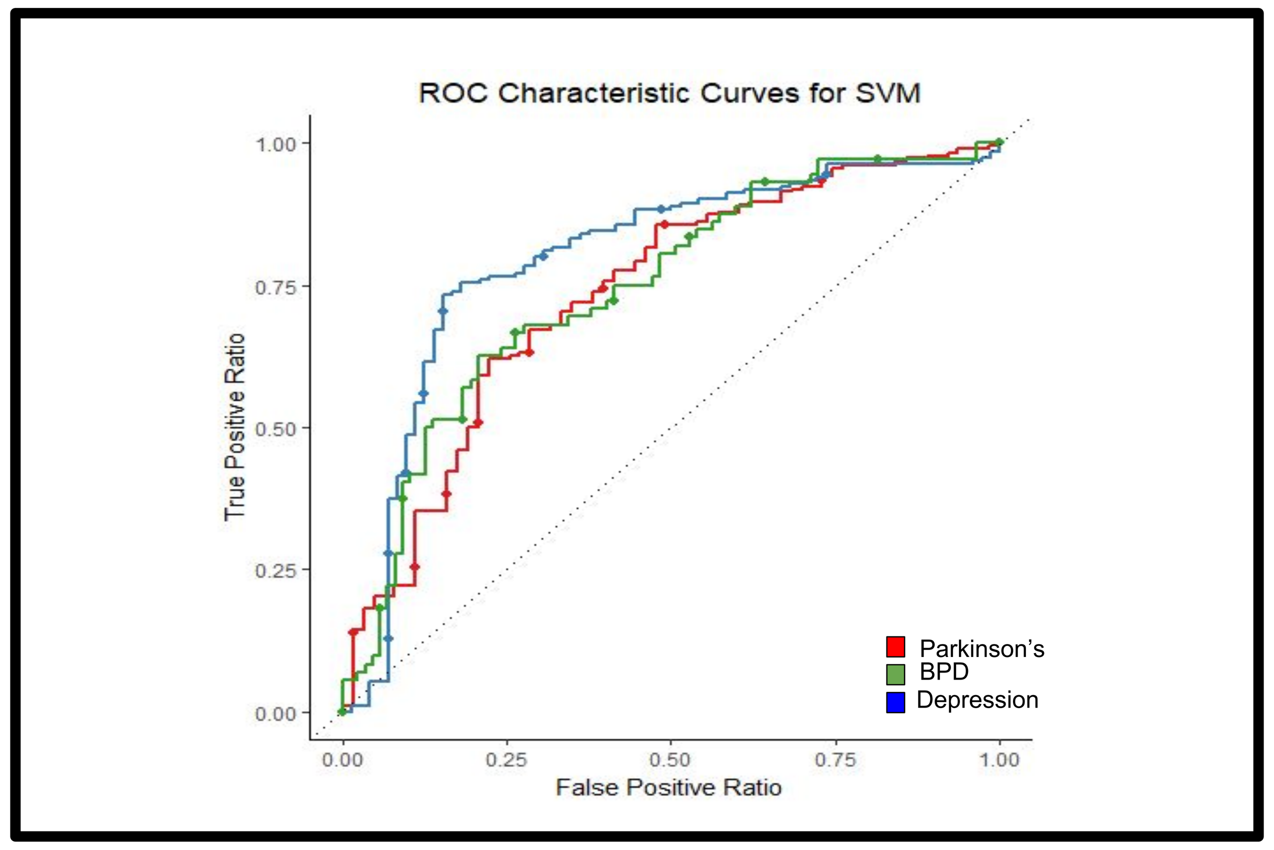

Figure 2: **Performance of Support Vector Machine (SVM) model on disease detection through voice.** Dotted line indicates 50% line ratio. Curves higher on the y-axis represent greater positive accuracy (i.e. sensitivity and specificity for disease presence).

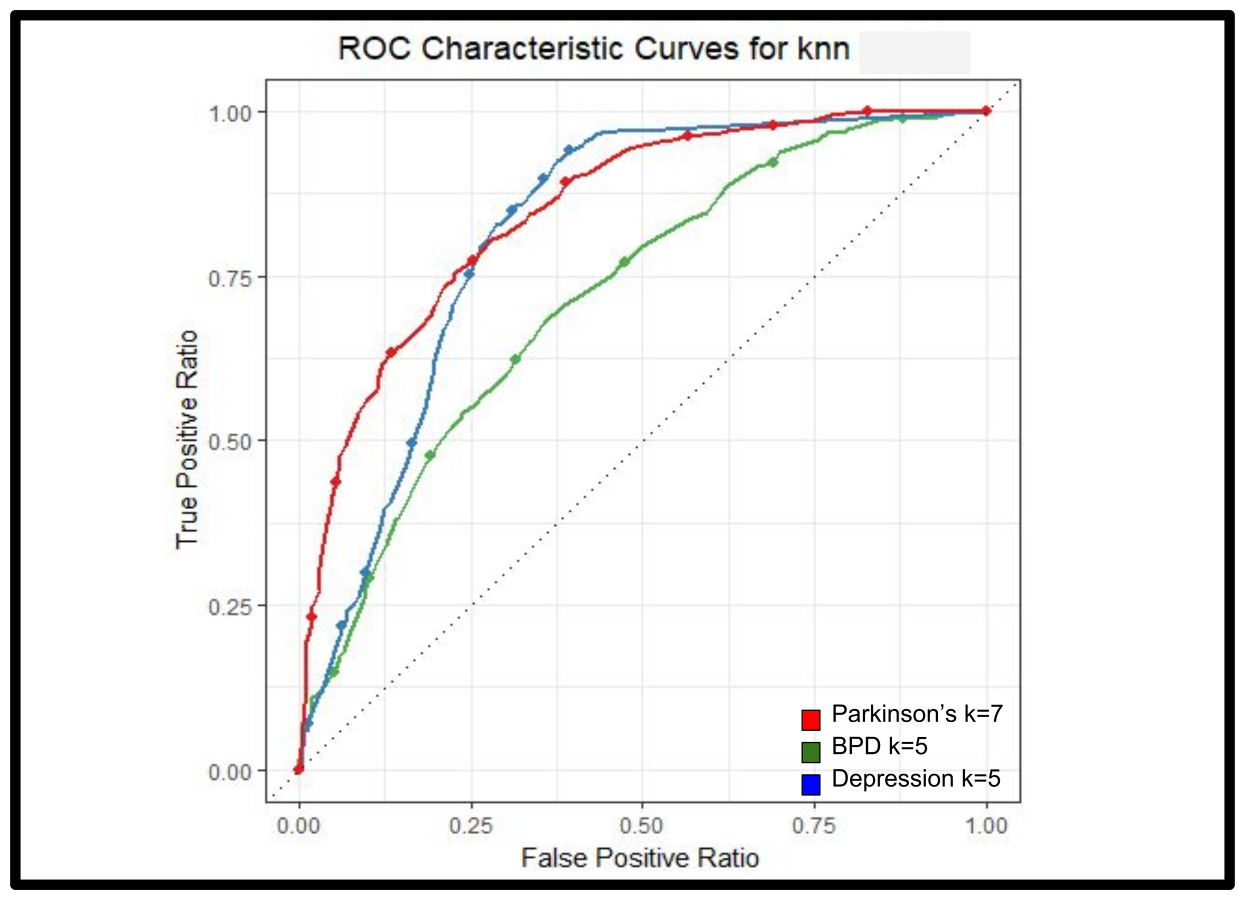

Figure 3: **Performance of k-Nearest Neighbors (kNN) model on disease detection through voice.** Dotted line indicates 50% line ratio. Curves higher on the y-axis represent greater positive accuracy (i.e. sensitivity and specificity for disease presence).

Results

- **Parkinson's Disease**
  - kNN was most accurate for PD
  - It was well predicted by RF, with an accuracy of over 90%, which is similar to the kNN model
  - The SVM strategy performed the lowest prediction of the three models with 80% accuracy
- **Depression**
  - RFwas most accurate for Depression
  - All three strategies performed fairly similarly, with a 80-85% accuracy
- **BPD**
  - Best predicted by RF, with a near 95% accuracy
  - kNN was able to predict with around 85% accuracy, and SVM was able to predict with around 75% accuracy
- Curves represent accuracy of the strategy on different diseases
  - Sensitivity and specificity were both factors that affect accuracy

Conclusion

- The three computational strategy models show potential to indicate presence of PD, Depression, and BPD utilizing same voice features over multiple models
- In addition, we will want to conduct further studies with more uniform sampling. This can be achieved through:
  - Larger sample size
  - More controlled settings (i.e. background noise, standardized voice collection)
  - Gender, ethnicities, exploration
  - More detectable differences in strategies for each illness
- Benefits:
  - Inexpensive, effective, non-invasive, and quick
  - Provides promising early diagnostic tool for health care through quantitative measure
  - Accessibility and portability
- Limitations: follow-ups with patients, survey accessibility, and recruiting patients
- Future Research Aims:
  - Distinguishing between diseases
  - Diagnosing disease severity
- Feature engineering to be incorporated in future to select features that perform the best

References

Holmes, R, et al., int. j. lang. comm. Dis. 35: 407–418 (2000) **[1]** I Yang, Y et al., IEEE Trans Affect Comput. 2:142-150 (2013) **[2]** I Teixeira AL, et al., Expert Opin Med Diagn. 7:147-59 (2013) **[3]** I Tysnes, OB, et al., J. Neural T 124: 901-905 (2017) **[4]** I Rizek, P, et al. Diag & T of Park 188: 1157-1165 (2016) **[5]** I NIMH, Major Depression **[6]** I Smith, K.M, et al. Comp. Psych. 54: 1-6 (2013) **[7]** I NIMH, Personality Disorder **[8]** I NIMH, Borderline Personality Disorder **[9]**

Acknowledgements

Data was contributed by patients from the University of Washington clinics and WA State area. We thank them for their support and trust. This work would not be possible without them. This study was reviewed by UW IRB (study 0790).

Dr. Hosseini Ghomi's work was supported by NIH R25 MH104159.

Additional support by Aaron Ong, Ali Rafiq, Bowen Xiao, Katina Papadakis, Kaylani Tam, Kielan Lemoine-Kowalski, Michelle Gouw, Sarah Holden, Prithvi Shetty, Larry Zhang