NAME: Dillon Pruett, dillon.g.pruett@vanderbilt.edu
PI: Jones, Robin
First theme choice: Clinical/Behavioral/Intervention

Creating sex-specific phenome risk classifiers via clinical comorbidities to identify under-documented cases of developmental stuttering in electronic health records

Authors: Dillon G. Pruett, Douglas M. Shaw, Hung-Hsin Chen, Lauren E. Petty, Hannah G. Polikowsky, Shelly Jo Kraft, Robin M. Jones, and Jennifer E. Below

Introduction: Though heritability estimates vary across studies, there is clear evidence that a genetic component for stuttering exists. Additionally, stuttering is more common in males, with a male-to-female ratio approximately 2:1 in children under 4 and 5:1 in adolescents and adults, and clinical evidence suggests a high incidence of comorbid speech, language, and attention disorders. Our study aims to examine potential sex-specific stuttering comorbidities and sex-specific genetic variants associated with stuttering by utilizing electronic health records (EHRs) paired with DNA biobank samples to perform 1) a sex-specific comorbidity analysis and, 2) a sex-stratified genome wide association study (GWAS).

Methods: First, individuals who stutter within the EHR will be identified using a keyword search, a text-mining algorithm, and manual review. Confirmed cases will then be stratified by sex and matched to controls in a permutation-based, sex-stratified comorbidity analysis of phecodes, hierarchical diagnostic groupings for EHR data derived from ICD-9 codes. Using the enriched phecodes identified from the comorbidity analysis, a phenome risk classifier will be developed and used to identify stuttering cases in Vanderbilt University Medical Center's DNA biobank (BioVU). Using the enriched phecodes, a phenome risk classifier will be developed and used to identify stuttering cases in BioVU. Finally, a sex-stratified GWAS will be conducted to identify variants associated with stuttering that differ between the sexes, which may reveal potential sex-specific gene involvement in this common, heritable, and often impactful disorder.

Results: Sex-stratified enrichments may replicate previous preliminary EHR-based comorbidity findings or reveal novel, sex-dependent comorbidities. Preliminary results from a smaller, non-sex-stratified subset of 572 stuttering cases found 38 phecodes significantly associated with stuttering. The proposed project seeks to replicate and extend these findings by examining potential sex-specific comorbidities in a larger dataset. Additionally, significantly enriched phecodes will be used to create sex-specific phenome risk classifier prediction models to identify additional high-likelihood stuttering cases that lack clinical documentation. A previous study created and tested a non-sex specific phenome risk classifier with a positive predictive value of 83%, validating this approach for acquiring developmental stuttering cases (Pruett et al., 2021).

Discussion: Greater understanding of stuttering comorbidities, especially sex-specific comorbidities in this highly sex-skewed population, has the potential to impact clinical care management and enhance patient care through advancements in identification, intervention, and treatment of stuttering. Furthermore, the creation of the sex-specific phenome risk classifiers will increase the number of high-likelihood stuttering cases identified within BioVU to conduct well-powered sex-specific genome-wide association studies to investigate the genetic etiology of this highly heritable and sex-skewed disorder.

References, if any: Pruett, D. G., Shaw, D. M., Chen, H. H., Petty, L. E., Polikowsky, H. G., Kraft, S. J., ... & Below, J. E. (2021). Identifying developmental stuttering and associated comorbidities in electronic health records and creating a phenome risk classifier. Journal of fluency disorders, 68, 105847.

Keywords:
Stuttering, Electronic health records, Comorbidity