Electrical Phenotypes of Genetic Risk for Atrial Fibrillation: Proposal for Validation Cohort

AF is a multifactorial disease that is both highly prevalent and burdensome, with a lifetime incidence of 1 out of 3-4 individuals. Although We have a limited understanding of the clinical factors that lead to atrial fibrillation (**AF**) pathogenesis, there is growing evidence of the genetic underpinnings of AF.1,2 It is clear that AF maintenance and pro-generation are a property of the atrial substrate, which can likely be further phenotyped based on how the atria are remodeled. The relationship of the atrial substrate to underlying genetic factors has somewhat been seen in studies evaluating electrocardiogram (ECG) associations with genetic variants.1,3–6 In this proposal, I plan to correlate the genetic manifestations of AF with the electrical properties of the atria through electrocardiographic (**ECG**) evaluation of P-wave morphology using a deep neural network (**DNN**).

Previously, as a feasibility study, I trained a DNN to predict the likelihood of the presence of a *TTNLOF* (loss-of-function) based off of 12-lead ECG tracings, with balanced accuracy of 82%. The model was trained on a cohort of n=200 individuals with known paroxysmal AF, with n=100 being reserved as a testing data set. Of those, using the strictest variant effect prediction definition for LOF, n=4 individuals were considered as cases (roughly 2% of the cohort). Each individual contributed 1 to 10 ECG, with each ECG being segmented into individual sinus beats, and each sinus beat used as a training (or test) data array. The findings were most prominent within the P-wave and PR interval. This model may have potential to be used to help screen for underlying genetic mutations that are related to AF, creating an *ECG-informed risk assessment*.

After confirming feasibility, I developed an expanded definition of cases and controls to extend the utility of the model. I separated known AF-related genes into two categories based on their putative protein: 1) ion channel proteins and 2) sarcomeric/structural proteins. I then compared these groups iteratively to a *wild-type* group, defined as individuals with no known deleterious AF-related genetic variants. Using this preliminary data, I aim to validate the model in a larger cohort of individuals with known AF and whole exome/genome sequencing (**WES/WGS**) data.

1. **Monogenic Aim**: Validate a DNN model to identify deleterious gene variants associated with AF in individuals with known AF, with cases defined as those with: a) sarcomeric or structural proteins, and b) ion channel-related proteins. The DNN would compare cases as those with gene variants, and use individuals with no known associated variants as controls. The proposed architecture would be similar to that of,7 which validates a similar DNN model on a monogenic arrhythmia/cardiomyopathy gene (phospholambdan).
2. **Polygenic Aim**: Train a DNN model to assess the polygenic risk (PGS) of AF in individuals with known AF and dense SNP genotyping data, based on recent, validated AF-GWAS polygenic risk scores.8,9 The model would be trained to predict the percent PGS of AF, similar to a linear regression, however the input is a 12-lead ECG.

I utilized participants from the UIC Multi-Ethnic AF Registry as the primary study population, who have a combination of WES and dense SNP genotyping data. This allowed me to generate a small DNN model for both monogenic and polygenic patterns. However, the sample size was small, reducing generalizability.

I propose to utilize a large, external cohort, such the UK Biobank (n=68,091 with complete data), as well as additional populations such as the ARIC cohorts, the US-CHARGE participating studies, and/or the Cardiac Heart Study. The All of Us study is limited to only 30% of the population with ECG tracings. I propose testing (and/or additional training) on these large cohorts to validate the model for monogenic and polygenic ECG-phenotypes of AF. My approach is the next logical step in better understanding the relationship between genetic variants and AF risk over time, and helps to create an *ECG-informed risk assessment* for the phenotypes of AF.

1. Weng L-C, Hall AW, Choi SH, Jurgens SJ, Haessler J, Bihlmeyer NA, et al. Genetic Determinants of Electrocardiographic P-Wave Duration and Relation to Atrial Fibrillation. Circ Genom Precis Med 2020;13:387–395. doi:[10.1161/CIRCGEN.119.002874](https://doi.org/10.1161/CIRCGEN.119.002874).

2. Choi SH, Jurgens SJ, Weng L-C, Pirruccello JP, Roselli C, Chaffin M, et al. Monogenic and Polygenic Contributions to Atrial Fibrillation Risk. Circulation Research 2020;126:200–209. doi:[10.1161/CIRCRESAHA.119.315686](https://doi.org/10.1161/CIRCRESAHA.119.315686).

3. Verweij N, Benjamins J-W, Morley MP, Van De Vegte YJ, Teumer A, Trenkwalder T, et al. The Genetic Makeup of the Electrocardiogram. Cell Systems 2020;11:229–238.e5. doi:[10.1016/j.cels.2020.08.005](https://doi.org/10.1016/j.cels.2020.08.005).

4. Holmqvist F, Kim S, Steinberg BA, Reiffel JA, Mahaffey KW, Gersh BJ, et al. Heart rate is associated with progression of atrial fibrillation, independent of rhythm. Heart 2015;101:894–899. doi:[10.1136/heartjnl-2014-307043](https://doi.org/10.1136/heartjnl-2014-307043).

5. Husser D, Stridh M, Sörnmo L, Roden DM, Darbar D, Bollmann A. A Genotype-Dependent Intermediate ECG Phenotype in Patients With Persistent Lone Atrial Fibrillation. Circulation: Arrhythmia and Electrophysiology 2009;2:24–28. doi:[10.1161/CIRCEP.108.799098](https://doi.org/10.1161/CIRCEP.108.799098).

6. Choi SH, Jurgens SJ, Haggerty CM, Hall AW, Halford JL, Morrill VN, et al. Rare Coding Variants Associated With Electrocardiographic Intervals Identify Monogenic Arrhythmia Susceptibility Genes. Circ Genom Precis Med 2021;14:e003300. doi:[10.1161/CIRCGEN.120.003300](https://doi.org/10.1161/CIRCGEN.120.003300).

7. van de Leur RR, Taha K, Bos MN, van der Heijden JF, Gupta D, Cramer MJ, et al. Discovering and Visualizing Disease-Specific Electrocardiogram Features Using Deep Learning. Circulation: Arrhythmia and Electrophysiology 2021;14:e009056. doi:[10.1161/CIRCEP.120.009056](https://doi.org/10.1161/CIRCEP.120.009056).

8. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018;50:1219–1224. doi:[10.1038/s41588-018-0183-z](https://doi.org/10.1038/s41588-018-0183-z).

9. Marston NA, Garfinkel AC, Kamanu FK, Melloni GM, Roselli C, Jarolim P, et al. A polygenic risk score predicts atrial fibrillation in cardiovascular disease. European Heart Journal 2023;44:221–231. doi:[10.1093/eurheartj/ehac460](https://doi.org/10.1093/eurheartj/ehac460).