Phenotyping AF

Project Page

Anish S. Shah, MD|MS

Background

Paroxysmal AF contains a broad spectrum of disease, regardless of burden, frequency, hemodynamic or vascular complications, or etiology. Several attempts have been made at pursuing more salient models, however they combine both paroxysmal and persistent AF into a binary diagnosis. Two recent studies have documented 4 potential clusters may exist (Pastori et al. 2020; Vitolo et al. 2021):

1. Younger, lower comorbidities
2. High CV risk factors
3. High CV comorbidities
4. High rates of non-CV comorbidities (e.g. cancer)

Wineinger et al. (2019) showed that in terms of duration and episodes of AF, there may be a *staccato* (short, frequent episodes) versus *legato* (long, infrequent) patterns to AF, however unknown about underlying clinical relevance.

In this proposal, I aim to phenotype paroxysmal AF into clinically-relevant subtypes. This serves as the basis for future work for evaluating differences in prognosis (including endocardial mapping patterns to recurrence).

Hypothesis

Hypothesis: *Diagnoses of paroxysmal AF will segregate into clusters based on electrical, clinical, and structural cardiac properties.*

The overall approach will be to use high-dimensionality reduction techniques to create finite clusters that alllow for segregation based on mutliple vector-spaces. This approach also allows for future *multi-state modeling* and *survival* techniques based on "events" → ablation, cardioversion, medication intensification, changes in burden, cardiomyopathy, etc.

Data

Data sources

The data that is currently available to answer this is derived from three sources:

a. UIC Multi-Ethnic AF Registry: ~1800 patients with various classes of AF that have contributed blood sampling for genetic analysis
b. Clinical Data Warehouse: collected by the CCTS, houses Cerner and EPIC charts from 2010 onwards, including clinical notes, procedure notes, diagnoses, vitals, medications
c. MUSE: ECGs from ~2000s to today with full digital/raw signal (10 seconds x 8 leads)

Key variables

The key variables that will be used are a combination of clinical, electrical, imaging that will allow for creation of multiple dimensions for analysis.

Electrocardiography

Sinus

- P-wave indices: voltage, AUC, duration, dispersion, terminal forces
- Global parameters: PR interval, P/R amplitude ratio
- Rate: high-frequency HRV, HR

AF

- Conduction properties: ventricular response, shortest/longest RR intervals
- Atrial parameters: dominant frequency (Fourier multi-transform), F-wave relative voltage + voltage over time

Clinical

- Age, Sex, Race
- Weight + weight over time
- Comorbidities (HTN, DM, HLD, CHF, CAD)
- CHADS2VAS2C score
- ADHF and h/o CHF
- Anti-arrhythmic intensification (AAD, BB uptitration, cardioversion, ablation)
- Atrial arrhythmias + recurrence + burden of AF over time
- Blood pressure + over time
- HR and responsiveness of HR over time

Echocardiography

- Volumes: LA, RA, LV, LVIDD
- Pressures: RVSP, PA pressures
- Diastology: E/e' ratio, A wave duration/velocity, A wave reversal
- Function: LVEF, RVEF, LA strain

Methods and approach

There are four components of this work that will eventually come-together for analysis.

a. Electrocardiography-based rhythm and morphology evaluation: capitalizes on arrhythmia-focused research while utilizing signal processing and ML-based clustering approaches
b. Echocardiography and structural imaging assessments of cardiac function: low-level NLP in structured data parsing
c. Clinical context including co-morbidities, vitals, medications, procedures: epidemiology-focused contextualization of patient attributes/trajectories
d. Changes in parameters with time: survival analyses and recurrent event analyses (multi-state models)

The initial focus will be on ECG-based signal processing approaches. This will involve downloading/extracting MUSE data in an XML format, converting the binary data into digital, and finally converting to the WFDB format (Waveform Database Software Package). Then, the WFDB binary files can be annotated for QRS location, beat segmentation, P-wave evaluation, etc. Using ML approaches, can subsequently convert and derive intervals, identify/confirm rhythm, and generate related parameters such as vector gradients. This will be used to generate features both in AF and in sinus rhythm. Using cluster-based approaches, can then identify if patterns emerge in ECG subtypes of paroxysmal AF.

Timeline
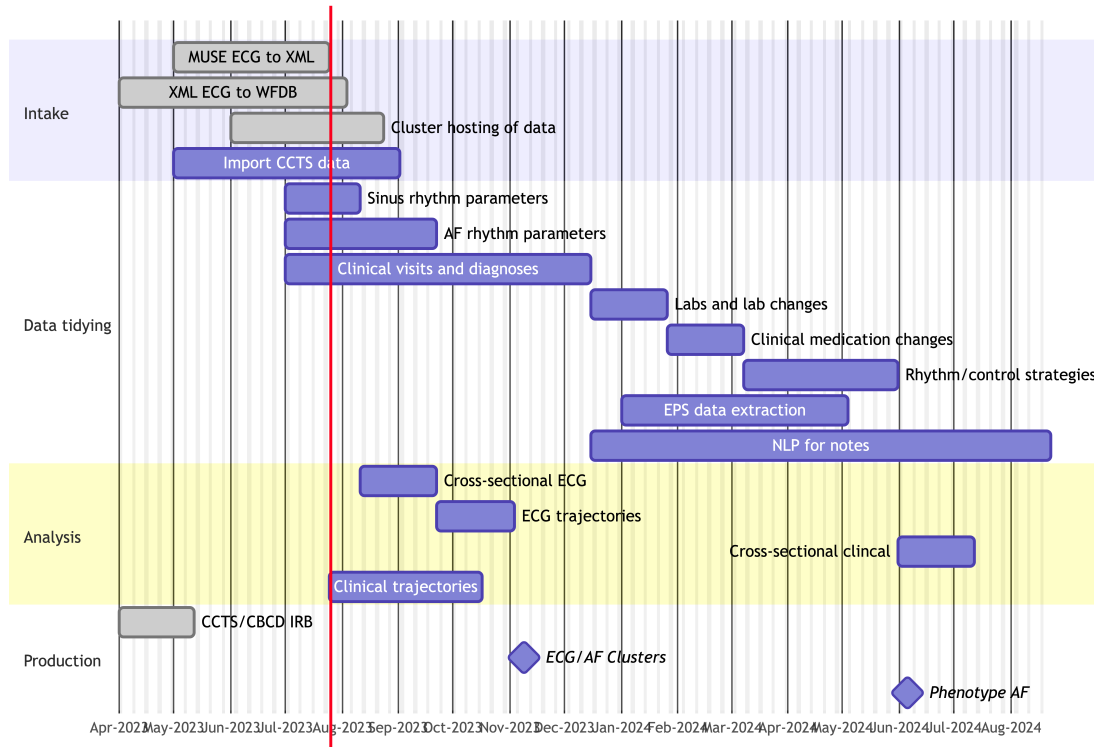
Pastori, D, E Antonucci, A Milanese, F Violi, P Pignatelli, G Palareti, and A Farcomeni. 2020. "Clinical Phenotypes of Atrial Fibrillation and Risk of Mortality: A Cluster Analysis." *European Heart Journal* 41 (November). https://doi.org/10.1093/ehjci/ehaa946.2893.

Vitolo, Marco, Marco Proietti, Alena Shantsila, Giuseppe Boriani, and Gregory Y. H. Lip. 2021. "Clinical Phenotype Classification of Atrial Fibrillation Patients Using Cluster Analysis and Associations with Trial-Adjudicated Outcomes." *Biomedicines* 9 (7): 843. https://doi.org/10.3390/biomedicines9070843.

Wineinger, Nathan E., Paddy M. Barrett, Yunyue Zhang, Ikram Irfanullah, Evan D. Muse, Steven R. Steinhubl, and Eric J. Topol. 2019. "Identification of Paroxysmal Atrial Fibrillation Subtypes in over 13,000 Individuals." *Heart Rhythm* 16 (1): 26–30. https://doi.org/10.1016/j.hrthm.2018.08.012.