CHOC/Mi-1 Research Proposal

**Background:**

Enola is an analytic framework derived primarily from Orphanet data which takes phenotypic and genotypic findings and returns a list of the most likely rare diseases, with the evidence for and against each disease. A web app has been developed which allows clinicians to manually input a patient’s findings and interact with the results.

**Study Aims:**

The first phase of this project aims to build a backend API that can communicate with the EHR through SMART on FHIR protocols to scan across a patient’s data and automatically extract findings, which are subsequently fed into Enola.

The second phase of this project aims to apply the Enola backend API to scan the EHR data for all CHOC patients who have a known rare disease and measure how frequently Enola places the correct diagnosis in the top 10 differential diagnoses.

The third phase of this project aims to apply the Enola backend API to scan across 2000 randomly selected patient records and identify the 100 patients who are most likely to have rare diseases, which is the approximate number we would expect to have real diagnoses given that approximately [5% of the population has a rare disease](https://www.nature.com/articles/s41431-019-0508-0). These 100 patient cases will be reviewed in I-Clinic to determine the number of children who have one of the diseases in the differential produced by Enola. A primary outcome is the number of patients who receive a new diagnosis of rare disease as a direct result of Enola’s investigations.

An ancillary objective of this study is to capture statistical associations among clinical entities in medical data, such as drug-disease associations. This will enable Enola to extend functionality beyond rare diseases to common diseases.

**Patient data handling and confidentiality**

* The application which scans patient data for findings will reside on CHOC servers or CHOC’s AWS instance, as will the application which learns statistical associations among clinical entities from the patient data.
* The API which takes a list of patient findings and returns the differential diagnosis can reside on any servers MI-1 designates, since no patient identifiers will be transmitted.
* The statistical associations among clinical entities learned during this project will be saved to an MI-1 database on any servers MI-1 designates, since no patient-level data will be transmitted.

**Key Stakeholders:**

* Sharief Taraman (PI for this project)
* Neda Zadeh (Project Advisor)
* Bill Feaster (CHOC VP & CHIO)
* Anthony Chang (Director at MI-1 and Chief Intelligence and Innovation Officer at CHOC)
* Brendan Dunphy (CEO at MI-1)
* Tim McLerran (Head of Product at MI-1)
* Potentially other CHOC-based data scientists, clinicians, etc

**Estimated resources required:**

* SMART on FHIR access to resources including: Medication.Search, MedicationRequest.Search, ServiceRequest.Read, DocumentReference.Search, DiagnosticReport.Search, Observation.Search, Procedure.Search, Condition.Search, Practitioner.Read, PractitionerRole.Read, AllergyIntolerance.Search, and others as may be identified during further development.
* Access to a sufficient number of patient records to validate the tool
* Business Associate Agreement to allow HIPAA-compliant access
* Technical performance test data
* Development to run 3 months starting in April 2022

**CHOC benefits:**

* Show publicly their ability to implement AI to enhance clinical practice
* Decrease the time to diagnosis for rare diseases
* Catch missed rare diseases, especially things CHOC hasn’t seen before