COVID-19 Evidence Accelerator Collaborative
Lab Meeting #4
Thursday, May 7, 2020, 3:00-4:00 pm ET

Call Summary

Update on Parallel Analysis Workstream

- The parallel analysis workstream has been working to establish a framework for answering key COVID-19 questions using real-world data (RWD).
- At this time, parallel analysis workstream participants are examining their individual data sets to provide further clarity around the process of collecting the data variables for analysis and simultaneously answering the same question across different datasets.
- Over the next two weeks, the workstream will be focused on:
    - Performing a descriptive analysis of COVID-19
    - Examining and describing the outcomes of patients receiving hydroxychloroquine (HCQ) and patients not receiving HCQ
- If your organization has data that could be contributed to this ongoing workstream, please reach out to Carla (RUF) or Jeff (FOCR).
- This workstream is also in the process of finalizing a statistical analysis plan (SAP) for COVID-19 research purposes. If your organization is interested in utilizing this SAP for its own research, please reach out to Carla (RUF) or Jeff (FOCR).

Introduction to Lab Meeting 4

The theme for this week's lab meeting was "building a machine for real-world data," and, more specifically, understanding the possibilities when we scale real-world research for COVID-19. Three presentations were given:

1. Presentation on the development of COVID-19 risk models & what they teach us about patient-level response to SARS-CoV-2 (Amy Justice, VA/Yale)
2. Presentation on common tools to support the design of COVID-19 clinical trials (Anne Heatherington, Takeda)
3. Presentation on a translation table for COVID-19 RWD common data elements and common data models (Mitra Rocca, FDA)

The Accelerator hopes to draw learnings from and build off these approaches in its parallel analysis workstream.

Lab Meeting Presentations

Presentation on the development of COVID-19 risk models & what they teach us about patient-level response to SARS-CoV-2

- This group has been tasked by FDA with answering a series of nine questions related to COVID-19. This presentation addressed the second question: What are the risk factors for mortality from COVID-19?
    - From a public health perspective, we need to know: who among us is most likely to die from COVID-19? This will help us take steps to prevent those most vulnerable from infection and will require data about pre-existing conditions.
    - From the clinical perspective, we need to know: what acute presentation suggests greatest risk of death? This will help us identify those most in need of medical treatment and will require data about pre-existing conditions and acute presentation.
- This group created a statistical model with the goals of:
    - Transparency—identify major risk factors and their impact on mortality in an understandable manner
    - Generalizability—maximize the likelihood that the model maintains its discrimination in new data
- They developed two data sets: a development set and a validation set consisting, in total, of over 5,000 patients.

Pre-existing Conditions Model

- They observed no association between race/ethnicity and mortality and a weak association which fell out after adjustment between gender, ACE/ARB/NSAIDS, body mass index, and specific comorbid conditions and mortality.
- They observed that the strongest predictor of mortality was age. While it is not uncommon to see age associated with mortality, the strength of the observed association is unique.
- This group utilized the Charlson Comorbidity Index to score patients using their comorbidities and observed that patients with a higher index had an increased % mortality. Multimorbidity is the salient risk factor, not specific comorbidities.
- Two surprising associations were observed:
    - Current and never smokers had similar risks of mortality, but former smokers had greater risk of mortality
    - Mortality seems to decrease as alcohol consumption increases in a stepwise manner
- After observing the association between smoking status and mortality, this group performed a gene expression comparison of bronchoalveolar lavage samples which demonstrated that many genes that are being downregulated in smokers are being upregulated in COVID-19 patients (and conversely, upregulated in smokers and downregulated in COVID-19 patients) which may lead to a better mechanistic understanding of the disease.
    - Smoking represses a key mechanism that we are studying in COVID-19 that explains most of the COVID-19 symptoms.
    - Chronic alcohol/smoking may dampen inflammatory response to COVID-19.

Acute Factors Model

- This group observed no association between pulse and mortality and a weak association, which fell out after adjustment, between fever and hemoglobin and mortality.

- While hypoxia, hypotension, white blood cell count and % lymphocytes were indepdently associated with mortality, acute kidney (eGFR) and liver (albumin, FIB-4) injury were more predictive.
- They observed that pre-existing model variables are reduced in significance once acute presentation is considered.
- Both models validated well in the prospective validation set.

Presentation on common tools to support the design of COVID-19 clinical trials

- The COVID R&D Consortium is a growing alliance of biopharma who wish to collaborate in work against COVID-19.
- The COVID R&D data-sharing workstream is focused on two main aspirations related to COVID-19:
    - Sharing emerging clinical trial data in a centralized repository
    - Maximizing the use of RWD to inform clinical trial design and accelerate development and/or registration
- In order to achieve these aims, the COVID R&D Consortium recognizes that, across the health care ecosystem, definitions, methods, data structure, algorithms, and endpoints need to be standardized. This standardization will better allow cross study comparison.
- Under the clinical trial aim, the COVID R&D Consortium is:
    - Seeking commitments on unprecedented sharing of clinical trial data and key documents across biopharma with permission for re-use.
    - Recognizing the need for infrastructure to support proposed data sharing, including secure platforms with access control.
- Under the RWD aim, the COVID R&D Consortium is working to:
    - Establish clear definitions and align on coding to make work consistent and reproducible across members
    - Understand different datasets, work with providers to enrich the data and ensure data are fit-for-purpose
    - Agree on key questions that facilitate clinical trials
    - Understand the scope of use (e.g., patient selection, inform endpoints, enhance control arms, facilitate findings)
    - Share and discuss outputs
- The R&D Consortium understands that this collaboration can lead to greater utility of all data sources: RWD can be used to inform and support the design of clinical trials and clinical trial data can be used to confirm RWD findings.

Presentation on a harmonization process from the COVID-19 RWD common data elements to several common data models and data standards

- The Common Data Model Harmonization Project has been underway for three years with a goal of building data infrastructure for conducting research using RWD derived from the delivery of health care in routine clinical settings.
    - Objective: Develop the method to harmonize the Common Data Models of various networks, allowing researchers to simply ask research questions on much larger

amounts of RWD than currently possible, and leveraging open standard and controlled terminologies to advance PCOR.

- Phase 1 accomplishments of this effort include:
    - Harmonized 5 common data models (i.e. Sentinel, PCORnet versions 3.1 and 4.0, OMOP, and i2b2/ACT) with an intermediary model (Biomedical Research Integrated Domain Group (BRIDG))
    - Developed the infrastructure (in collaboration with the NIH/NCATS) to build a query, view and store the results leveraging open, consensus-based standards
    - Collaborated with Yale/Mayo Clinic as well as Elligo Health Research on the execution of the query focusing on the safety of immuno-oncology products in patients with autoimmune disorders.
- Planned Phase 2 deliverables include:
    - Collaborate with new data partners leveraging the CDMH architecture as well as direct query from electronic health records and clinical data repositories.
    - Enhance the existing infrastructure to leverage Health Level Seven (HL7) Fast Healthcare Interoperability Resources (FHIR) standard as the exchange data standard.
    - Submit RWD leveraging clinical trial study data, leveraging Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM) via the FDA Gateway
- The National COVID-19 Cohort Collaborative (N3C) is:
    - A centralized, secure portal for hosting row level COVID-19 clinical data and deploying and evaluating methods and tools for clinicians, researchers, and healthcare
    - A partnership among several HHS agencies, the CTSA network, distributed clinical data networks (e.g. PCORnet, OHDSI, ACT/i2b2, and TriNetX), and other clinical partners
    - Four community workstreams:
        - Data Partnership & Governance
        - Phenotype & Data Acquisition
        - Data Ingestion & Harmonization
        - Collaborative Analytics
- The list of COVID-19 data elements identified by FDA, RUF and Friends of Cancer has been mapped to several Common Data Models (CDMs) and standards. Steps in this process are:
        - Review and augment the COVID-19 data elements
        - Map COVID-19 data elements to PCORnet, i2b2/ACT, Sentinel, OMOP CDMs
        - Map the COVID-19 data elements to the United States Core Data for Interoperability (USCDI) and Health Leven Seven (HL7) Fast Healthcare Interoperability Resources (FHIR) R4
        - Map the COVID-19 data elements to Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM) Standard and COVID-19 companion guide
        - Map the COVID-19 data elements to VA EHR system Data Elements
        - Validate the mappings with the Standards Development Organizations (SDOs) and the technical leads for each CDM

- The NIH Common Data Elements (CDE) Task Force seeks to identify CDEs for COVID-19 studies, EHR data, and other data sources and is interested in exploring the use of these mappings.
- The I-SPY COVID Trial is an adaptive platform trial designed to efficiently and effectively find agents with the most potential to reduce mortality/morbidity. The I-SPY COVID Trial will leverage the OneSource platform developed for I-SPY 2 Trials focusing on the use of EHRs in breast cancer trials.