



## **COVID-19 Evidence Accelerator Collaborative**

### **Lab Meeting #7**

*Thursday, May 28, 2020, 3:00-4:00 pm ET*

### **Call Summary**

#### **Introduction to Lab Meeting 7**

The theme for this week's lab meeting was "Natural History Gets Serious: Scalable Approaches to Define and Document Natural History." As we work towards solving critical COVID-19 questions, a solid understanding of the natural history of the disease will optimize the design of clinical trials, enable benchmarking of patient and disease status, aid in the development of prediction models, provide insights into appropriate patient care, and inform planning for vaccines.

Professionals involved in the Observational Health Data Sciences and Informatics (OHDSI) collaborative gave three presentations and live demonstrations of their current work investigating the natural history of COVID-19:

1. Presentation on An International Characterization of Patients Hospitalized with COVID-19 and a Comparison with Those Previously Hospitalized with Influenza (Dani Prieto-Alhambra, University of Oxford; Patrick Ryan, OHDSI)
2. Presentation on Seek COVER: Development and Validation of a Personalized Risk Calculator for COVID-19 Outcomes in an International Network (Ross D. Williams, Erasmus University Medical Center; Patrick Ryan, OHDSI)
3. Presentation on Characterizing Health Associated Risks and Your Baseline Disease in SARS-CoV-2 (CHARYBDIS) (Talita Duarte-Salles, IDIAPJGol; Primary Care Research Institute Jordi Gol) Patrick Ryan, OHDSI)

Before the lab meeting presentations, Patrick Ryan gave an overview of the COVID-19 work OHDSI is doing across three different use cases: clinical characterization, patient-level prediction, and population-level effect estimation. He also emphasized the importance of being fully transparent in every step of the research process, a theme which was reiterated throughout the course of the meeting and is demonstrated by OHDSI's sharing of its research protocols, definitions, results, and analyses in the public domain. As OHDSI states, their studies are "living evidence repositories" and "data partners can execute analyses and share aggregate results at any point, including updates as data accumulate."

#### **Lab Meeting Presentations**

##### **Presentation on An International Characterization of Patients Hospitalized with COVID-19 and a Comparison with Those Previously Hospitalized with Influenza**

- This study aggregated data representing 26,074 patients hospitalized with COVID-19 in the United States (2,477 patients), South Korea (5,172 patients), and Spain (18,425 patients).

- From these patients, 49,331 summary characteristics were extracted and summarized in an interactive web-based application.
- Key observations include:

- Patients were majority male in the US and Spain, but majority female in South Korea.
- Age profiles varied across data sources.
- When compared to patients hospitalized with influenza, patients admitted to the hospital with COVID-19 tended to be healthier: they typically had fewer comorbidities and less medical drug usage. Exceptions to this trend were observed for specific comorbidities including obesity, diabetes, and osteoarthritis.
- OHDSI's sharing of aggregate study results in a web-based application in addition to its summarized key findings in a medRxiv publication is additive to the body of work being generated about COVID-19 natural history and demonstrative of the fully transparent and collaborative research process needed to rapidly answer COVID-19 questions.<sup>1</sup>
- Other individuals and organizations are encouraged to use OHDSI's online tool for answering additional questions of interest.

### **Presentation on Seek COVER: Development and Validation of a Personalized Risk Calculator for COVID-19 Outcomes in an International Network**

- This study developed a model for predicting COVID-19 outcomes in symptomatic patients presenting in the outpatient setting by building on work previously done in two publications:
  - "Design and Implementation of a Standardized Framework to Generate and Evaluate Patient-Level Prediction Models in Using Observational Healthcare Data"<sup>2</sup>
  - "Feasibility and Evaluation of a Large-Scale External Validation Approach for Patient-Level Prediction in an International Data Network: Validation of Models Predicting Stroke in Female Patients Newly Diagnosed with Atrial Fibrillation"<sup>3</sup>
- The model development used a patient population consisting of over 2 million patients with a general practice (GP), emergency room (ER), or outpatient (OP) visit with diagnosed influenza or flu-like symptoms prior to 2020. The model was validated on patients with a GP, ER, or OP visit in 2020 with a confirmed or suspected COVID-19 diagnosis. Patients in both cohorts required 365 days of continuous observation prior to cohort start date (visit to GP/ER/OP).
- Pre-index characteristics used as predictors included:
  - Age defined as year (cohort start date) – year of birth: 5-year strata
  - Sex
  - Concept-based characteristics:
    - Condition groups (SNOMED + descendants), ≥ 1 occurrence
    - Drug era groups (ATC/RxNorm + descendants), ≥ 1 day during the interval which overlaps with at least 1 drug era
  - Cohort-based characteristics:

---

<sup>1</sup> <https://www.medrxiv.org/content/10.1101/2020.04.22.20074336v1>

<sup>2</sup> Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR. Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *J Am Med Inform Assoc.* 2018;25(8):969-975. doi:10.1093/jamia/ocy032

<sup>3</sup> Reps JM, Williams RD, You SC, et al. Feasibility and evaluation of a large-scale external validation approach for patient-level prediction in an international data network: validation of models predicting stroke in female patients newly diagnosed with atrial fibrillation. *BMC Med Res Methodol.* 2020;20(1):102. Published 2020 May 6. doi:10.1186/s12874-020-00991-3

- Cancer, chronic obstructive pulmonary disease, diabetes, heart disease, hypertension, hyperlipidemia, and kidney disease
- Using this model, the study constructed risk predictions for three outcomes of interest:
  - Hospitalization with pneumonia
  - Hospitalization requiring intensive services (mechanical ventilation, tracheostomy, or ECMO) or death
  - Death in the 30 days after index date
- Across these three outcomes of interest, risk distributions by COVER score varied markedly.
- As demonstrated by this work, developing and validating predictive models allows us to recognize that variables are highly dependent on one another and to make risk calculations based on this dependent relationship.
- The OHDSI predictive model which was calibrated across the OHDSI network allows us to make predictions about which patients with COVID-19 are likely to need hospitalization and provides initial insights about incidence of disease.
- A digital version of the risk calculator and additional information about OHDSI’s model development and validation are available at <http://evidence.ohdsi.org/Covid19CoverPrediction>.

**Presentation on Characterizing Health Associated Risks and Your Baseline Disease in SARS-CoV-2 (CHARYBDIS)**

- The CHARYBDIS study has two primary aims:
  - To describe the baseline demographic, clinical characteristics, treatments, and outcomes of interest among individuals with COVID-19 overall and stratified by sex, age, and specific comorbidities.
  - To describe characteristics and outcomes of influenza patients between September 2017 and April 2018 compared to the COVID-19 population.
- This study is interested in understanding the COVID-19 patient trajectory through all stages of the disease: health pre-COVID-19 > present with symptoms > tested for COVID-19 > result obtained for COVID-19 test > hospitalization > hospitalization with intensive services > death. A mortality follow-up observation period is built in as well (30-d, 90-d, and 365-d risk window)
- Target cohorts are:
  - COVID-19: persons tested for SARS-CoV-2; persons tested positive for SARS-CoV-2; persons tested with a COVID-19 diagnosis record or a SARS-CoV-2 positive test; persons with a COVID-19 diagnosis or a SARS-CoV-2 positive test; persons hospitalized with a COVID-19 diagnosis record or a SARS-CoV-2 positive test; persons hospitalized and requiring intensive services with a COVID-19 diagnosis record or a SARS-CoV-2 positive test
  - Influenza: persons with influenza diagnosis or positive test 2017-2018; persons hospitalized with influenza diagnosis or positive test 2017-2018; persons hospitalized with influenza diagnosis or positive test requiring intensive services 2017-2018
- The study will be stratifying based on a list of factors generated using input from the broader community.
- As a part of this body of work, the OHDSI group in collaboration with the community has identified the “vocabulary” used to identify patients who have received SARS-CoV-2 testing. This

vocabulary includes 53 distinct codes used to identify testing and 40 different LOINC codes which capture the “flavors” of the testing.

- The OHDSI group is also emphasizing the need for understanding the diversity of measurement values and diagnostic source codes being used for COVID-19 across the community.
- In line with OHDSI’s prioritization of a transparent research process, all phenotypes are being made available to the public online as a part of the CHARYBDIS initiative. Opportunities for a collaborative, open phenotype library for the scientific community to use and learn from should be explored.
- Individuals or organizations with access to COVID-19 data that is mapped or data that they are planning to map who are interested in participating in this initiative should reach out to the OHDSI team.

### **Discussion**

- It is important for the science community to run these disease natural history studies across different datasets to ensure that we are capturing accurate observations and identifying patterns which are consistent and reproducible across data sources.
- Natural history studies will continue to inform our understanding of COVID-19 and influence patient care.
- Now more than ever, it is important to be open and transparent through all phases of the research process. We must be clear about what the scientific research standards are and strive to continue raising the bar.