

**Title: COHD-COVID:
Publicly Shared COVID-19
Statistics Mined from EHR
data**

INTRODUCTION

Sharing and utilization of clinical data from different institutions, hospitals, and nations can facilitate COVID-19 research since important characteristics of the patients are often found by examining the shared clinical data.

METHODS

- Concept Prevalence Analysis

$$P_H^C = \frac{|T_H^C|}{|T_H|}$$

where P_H^C is prevalence of concept C in cohort H, T_H^C is the set of unique visits of patients observed with concept C, and T_H is the set of unique visits of patients in the cohort H. We used three different cohort*: (1) COVID-19 patient cohort; (2) Influenza cohort; (3) General cohort.

- Concept prevalence ratio analysis, Concept co-occurrence analysis, and symptom prevalence analysis will be added to COHD-COVID API.

RESULTS

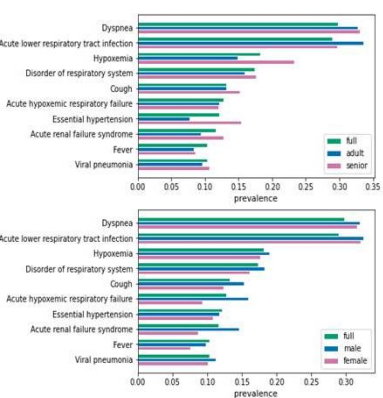


Figure 1. Condition concept prevalence in (a) age and (b) sex sub-cohorts of the COVID-19 cohort. The full COVID-19 cohort indicates original COVID-19 cohort without stratification.

*See our abstract for more details about the cohorts

COHD-COVID: Publicly shared COVID-19 statistics mined from EHR data

PRESENTER: Junghwan (Jay) Lee
Junghwan Lee, MA‡;
Casey Ta, PhD‡;
Jae Hyun Kim, PhD;
Cong Liu, PhD;
Chunhua Weng, PhD*

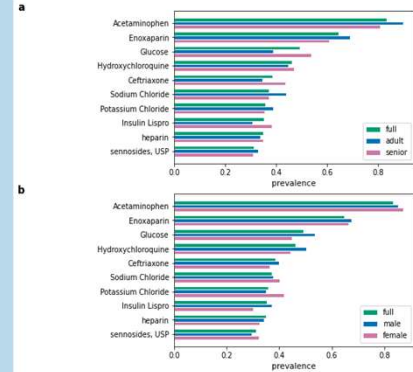


Figure 2. Drug concept prevalence in (A) age and (B) sex sub-cohorts of the COVID-19 cohort. The full COVID-19 cohort indicates original COVID-19 cohort without stratification.

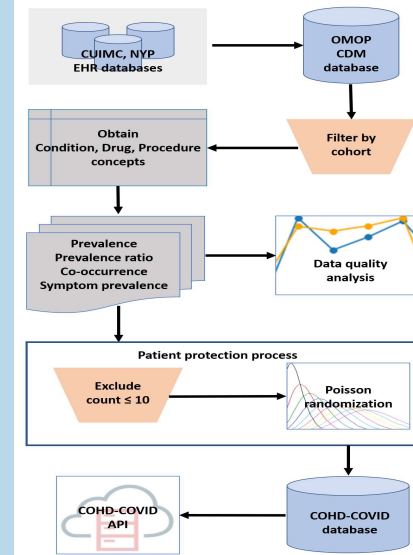


Figure 3. Overall workflow of COHD-COVID

	COVID-19 cohort	General cohort	Influenza cohort
Time range	2020.03.01 – 2020.05.31	2014.01.01 – 2019.12.31	2014.01.01 – 2019.12.31
# of patients	3,102	175,876	3,263
# of male patients (# of inpatient visits)	1,692	67,396	1,455
# of female patients (# of inpatient visits)	1,410	108,480	1,808
# of adult patients (# of inpatient visits)	1,445	106,729	1,371
# of senior patients (# of inpatient visits)	1,657	69,147	1,892