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
  \[ P_C = \frac{|T_C|}{|T_H|} \]
where \( P_C \) is prevalence of concept C in cohort H, \( T_C \) 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

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