Submission Id: 5404

Title

Harmonized Healthcare Database across Family Medicine Institutions

Priority 1 (Research Category)

Big Data

Presenters

Chance Strenth, PhD, David Schneider, MD, MSPH, Usha Sambamoorthi, Sravan Mattevada, MS, Kimberly Fulda, DrPH, Bhaskar Thakur, PhD, Masters of Population Studies, Anna Espinoza, MD

Abstract

Context: Electronic Health Record (EHR) data provides family medicine physicians with an accessible entry point into research on health outcomes of their patient. However, accessing EHR data can be a daunting process that can be stymied by institutional policies and regulations. Even when data is readily available, limitations on patient population and diversity may limit conclusions that can be drawn from it.

Objective: From the continuation of previous work, the purpose of this study is extended data harmonization efforts across three EHR institutions and two different EHR systems. Previous efforts included EHR data from a single calendar year, however, data harmonization now includes EHR data from six calendar years. The current study will describe this process and continuing challenges that have been faced.

Study Design: Cross-sectional and longitudinal study designs.

Setting or Dataset: Harmonized EHR data extracted from three primary care healthcare systems all from a large metropolitan area.

Population studied: All adult patients seeking care in family medicine clinics at three academic medical centers in North Texas between January 1st , 2016 to December 31st, 2022.

Outcome Measures: Prevalence of treated chronic medical conditions, clinical management, and biomarkers of chronic conditions.

Results: In total, data from 240,214 family medicine patients over the course of six years were extracted from the three EHR systems. The harmonized dataset consists of mostly women (N = 143,776), and similar sized groups of White (N = 70,439), African-American (N = 66,166), and Hispanic/Latino (N = 69,636).. The most common chronic conditions found were: arthritis, diabetes, hyperlipidemia, and hypertension. Data from the most frequent medications, and labs will also be presented. A discussion on challenges in harmonizing the different EHR systems such as issues concerning formatting, and missing values, and how they were addressed will also be presented.

Conclusion: Harmonizing EHR data allows for a more generalizable population to be analyzed and a streamlined process to address patient outcomes in family medicine. The current harmonized dataset contains data from a diverse metropolitan area and allows for the analysis of trends in disease progression. The current findings also provide a "roadmap" for other institutions to use similar methods to harmonize their EHR data.