Submission Id: 3552

Title

Using epidemiology and artificial intelligence to describe a complex primary care population in a learning health system

Priority 1 (Research Category)

Healthcare informatics

Presenters

Jacqueline Kueper, PhD, MSc, Jennifer Rayner, PhD, Merrick Zwarenstein, MD, PhD, Dan Lizotte

Abstract

Context: Electronic health records (EHR) provide an opportunity for developing decision support and other types of learning health system (LHS) initiatives. Careful understanding of the population of interest and how clients are represented in the data is essential for problem selection and for effective study design and analysis of data to solve the problem. The Alliance for Healthier Communities is one of the first primary care LHS in North America, serving complex, at-risk clients through Community Health Centres (CHCs) across Ontario, Canada. We propose that to properly understand their electronic health record data both simple statistical techniques commonly seen in descriptive epidemiology and more complex techniques from artificial intelligence will be useful.

Objective: To summarize characteristics of ongoing primary care clients served by CHCs for the purpose of informing future LHS initiatives.

Study Design and Analysis: Table-based summaries are given for all outcome measures. An Ising model is used to identify condition co-occurrences, non-negative matrix factorization is used to examine care provider teams, and time series clustering is used to explore care frequency patterns.

Setting or Dataset: EHR data from all CHCs across Ontario, Canada in 2009-2019.

Population Studied: Clients over 18 years old in 2009 who indicated a CHC as their primary care provider and had at least one encounter in 2009-2019.

Intervention/Instrument: N/A.

Outcome Measures: Sociodemographic characteristics (from structured client characteristic table, e.g., age, education), clinical characteristics (20 chronic conditions, four acute conditions, multimorbidity), and health care use characteristics (providers involved, complexity, frequency).

Results: There were 221,047 eligible clients. Social determinants and clinical conditions are most prominent in CHCs serving those most at risk in urban settings; the former are also more prevalent in clients with multimorbidity. Physician and nursing types provide most care, with heterogeneous combinations of other providers. There is notable within- and between-client variability in care complexity and frequency. We identify implications for future analyses such as cohort building and risk prediction.

Conclusions: Simple statistical and artificial intelligence techniques can be used to effectively describe a complex primary care population; this population-level overview provides a foundation for future LHS initiatives.