

Internal and External Data Linkage of Complex Relational Database: Results from CorHealth Ontario

Fang, J¹, Qiu, F¹, Wang, X¹, Fischer, H¹, and Tu, J¹¹Institute for Clinical Evaluative Sciences

Introduction

CorHealth Ontario, formerly Cardiac Care Network (CCN), maintains a registry of patients undergoing select cardiac procedures/surgeries in Ontario, Canada. This population-based database contains over 35 datasets with complex structure, linked by unique primary key or multiple keys.

Objectives and Approach

We aimed to simplify the complex CorHealth database so that research analysts could create study cohorts more efficiently and effectively, and to enrich the study cohort by getting more clinical information through database linkage. Through internal linkage, we could combine clinical fields from multiple CorHealth datasets. While the CorHealth dataset may not have all the clinical information needed for a given study, we may link the CorHealth study cohort externally to other administration databases to obtain additional fields via the probability matching (i.e., identical patient ID, hospital ID and procedure/surgery date).

Results

After identifying the primary keys on the relational database flowchart, we designed new data structures by combining similar topic datasets. The total number of datasets was reduced from 35 to 13. This simplified CorHealth dataset includes one main CorHealth dataset (including demographic information, referral data, comorbidities) plus 12 other linkable specific datasets (including stent, vessel, TAVI, STEMI). Through internal linkage, we can get the stent numbers, lengths and types of Percutaneous Coronary Interventions from the Stent dataset. Linking to Discharge Abstract Database (DAD), we can get the hospital length of stay and the episode of care of hospital transfer for each procedure; linking to The Ontario Health Insurance Plan database (OHIP), we can find the graft numbers and vessel types of Coronary Artery Bypass Graft.

Conclusion/Implications

To improve the research capacity and increase the value of the CorHealth database, analysts could create enhanced cardiovascular study cohorts derived from the simplified CorHealth database, plus internal linkage from other CorHealth datasets, and external data linkage from population-based administrative sources. We have accomplished three reports (PCI/CABG/TAVI) accordingly in 2017/18.

