Individuals and Organizations Completing Research in the IMEDS Lab
Vojtech Huser, MD, PhD, Assistant Clinical Investigator, Laboratory of Informatics Development, NIH Clinical Center

Parameters:
Start date: June 1, 2014 End date: Dec 1, 2015

Access request for: Vojtech Huser MD PhD, and John Kimbrough MD PhD (Collaborating post-doctoral student)

Description:
Comprehensive and complete electronic health record is an important assumption for many drug-safety and research analyses. This research project will focus on analyzing and measuring the comprehensiveness of Electronic Health Record (EHR) in a given Integrated Data Repository (IDR) (e.g., GE Healthcare dataset). We have previously developed an IDR Snapshot tool on two prior datasets (Marshfield Clinic and NIH Clinical Center) [1] that tries to measure data completeness. We plan to develop a more detailed measures looking at "average" or typical patients for a number of common health condition (eg hypertension, chronic kidney disease, heart failure). Our goal is to model a typical progression of the disease (taking into account majority of possible scenarios) and investigate whether a given dataset truly observes all phases of the process.

For example, a given dataset may lack outpatient data from a specific specialty and appear incomplete. This is because of underlying integrate data delivery network rather than due to a differences in the medical care provided. We hope to model using process mining a set of medical conditions and evaluate a given warehouse whether these scenarios are observed. For example, a complete record for chronic kidney disease patients [2] is important for any drug safety analyses conducted on those patients. By covering a diverse and large enough set of disease, we hope to capture overall completeness of a given dataset (eg, Truven Market Scan). We also want to model measures that will show that smaller overall data set can be superior to a larger dataset (for a given set of diseases) (for example, IDR with clinical trials data for a rare condition over a general population dataset).

We also plan to evaluate how new version of OMOP Common Data Model version 5 can better support data quality and process evaluation and what improvements it may need to better cover all important disease process aspects. We will work on CDM v5 features that could be used to capture data from NIH Clinical Trials data repository (called BTRIS) and pilot OMOP CDM v5 representation of BTRIS data. (Note: will work with the v5 draft specs only (for the time being) (not with v5 shaped data - since it will take time to create).

Supporting References:

1. Huser V, Cimino JJ. IDR Snapshot: Quantitative Assessment Methodology Evaluating Size and Comprehensiveness of an Integrated Data Repository. AMIA Translational Bioinformatics Summit 2012.
2. Huser V, Starren JB. EHR Data Pre-processing Facilitating Process Mining: an Application to Chronic Kidney Disease. AMIA Annu Symp Proc 2009.