Evaluating the Concordance between Patient Reported Data and Electronic Health Record Data: A Scalable Query Interface for Complex Analysis

Kimberly Barrett, MPH1; Jessica Malenfant, MPH1; Lesley Curtis, PhD2; Jill George, CCRA2; Sunil Kripalani, MD, M.Sc., SFHM3; Heather Limper, PhD3; Emily O’Brien, PhD, FAHA2; Tammy Reece, MS, PMP, CCRA2, Jeffrey S. Brown, PhD1

1Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA; 2 Duke University, Durham, NC; 3Vanderbilt University Medical Center, Nashville, TN

Introduction

The ADAPTABLE trial is the first major randomized comparative effectiveness trial to be conducted by the National Patient-Centered Clinical Research Network (PCORnet)1. ADAPTABLE encompasses several key features, including enrollment of 15,000 patients across 40 health care systems and health plans; a web portal for collecting patient-reported outcomes (PRO) data (e.g. medications and experiences); and reliance on existing EHR (Electronic Health Record) data for baseline characteristics and outcomes1. Longitudinal data on patient-reported information, not readily available in the EHR, is collected and provides the opportunity to evaluate the level of concordance between PRO and EHR data. The ADAPTABLE Supplement project was a unique opportunity to develop, pilot and evaluate methods to compare PRO data and EHR data stored in the PCORnet Common Data Model (CDM)2.

We will present our work that extends an existing Menu-Driven Query (MDQ) interface in the open source platform PopMedNetTM (PMN)3. MDQs enable investigators to compose and distribute custom, sophisticated queries via a user interface in PMN. We developed an analytic tool to quickly and efficiently compare patient-reported and clinical data. Enhancements to MDQs included expanding the querying capability to the Trial and Patient Reported Outcome tables within the PCORnet CDM, as well as developing temporal functionality that enables us to query for encounters that are within a specified time window of a patient-reported hospitalization date. All features developed for the ADAPTABLE Supplement study have been designed with the ability to re-use the tool for future studies.

Methods

The primary objective regarding MDQs was to develop a reusable tool that will output patient-reported information and clinical information regarding hospitalization dates and events side by side for concordance analysis. There were 3 use cases that defined our requirements. The project included one Data Partner, Vanderbilt University Medical Center (VUMC), who populated their PCORnet CDM in an Oracle database with approximately 2,000 patients with PRO data captured as part of the ADAPTABLE Trial. The MDQ tool relies on a point and click, flexible structure to identify cohorts of interest through defining preset Terms. The terms available are designed to query fields within the PCORnet CDM. New queryable Terms were added to the interface that target the tables describing trials (PCORNET_TRIALS) and patient-reported outcome measures and questionnaires (PRO_CM). The query output was extended and produces a line list of the patient-reported outcomes alongside associated encounter, diagnosis, and procedure code information. The specific functionality developed includes the following:

1. Ability to query the PCORNET_TRIALS table within the database. This enables the investigators to restrict the query to patients within the ADAPTABLE Trial.

2. Ability to query the PRO_CM table within the database. This enables the tool to query certain patient-reported outcome (PRO) fields within the database. This is key to extracting the PRO information of interest regarding hospitalizations.

3. Temporal event functionality. This provides the ability to enter an Index Event that is present within the PRO_CM table, as well as a desired time window. For example, a use case for the ADAPTABLE Supplement project included PRO Hospitalization date as the Index Event and the time window was set for 7 days before and 7 days after the index event. The MDQ then searches the database for Encounters that fall within the specified time window for each patient’s Index Event(s). The query will apply to every index event and will return in the output all encounters that match the criteria and time window.
The MDQ output has been expanded to include both the PRO and EHR information that results from the query, including information from the following tables: PCORNET_TRIAL, PRO_CM, ENCOUNTER, DIAGNOSIS, and PROCEDURES.

**Results**

The tool was tested and validated against synthetic data populated to match how the Data Partner would store their PRO data. Additionally, the ADAPTABLE Supplement team at VUMC successfully ran multiple queries of different levels of complexity to view the patient reported responses within the same row as the patient’s matching EHR information. The output in Figure 1 is for a query that was interested in finding any encounter that was within 1 day of the patient reported hospitalization date, plus information on the patient-reported hospitalization events. Figure 1 provides an example of what the output could look like for a participant in the ADAPTABLE Supplement study who reported two hospitalizations, both of which matched with an encounter in their EHR. The patient-reported events and dates are on separate rows due to how the PRO_CM table is populated. Investigators in the ADAPTABLE Supplement study were able to link the separate rows together through the use of the PRO_MEASURE_SEQ field (in circles in Figure 1). Several rows with the same number in the PRO_MEASURE_SEQ column indicates that they are all a part of the same hospitalization. Reading the first hospitalization in blue from left to right, we can see an Encounter on 9/1/2018 for the ICD-9 diagnosis of 434.91 (Stroke) and the patient reported event of a stroke on 9/1/2018. Due to matching dates and events, the patient reported hospitalization in blue is considered to be in concordance with that patient’s EHR information. The second hospitalization in red depicts an Encounter on 9/15/18 for a HCPCS procedure for a CABG (two veins). The patient reported undergoing a CAGB surgery on 9/14/18. In the case of the second hospitalization, the patient reported the correct procedure, but the date was one day before the date reported in the health record.

**Figure 1.** MDQ Output for Test Participant #W12.

**Discussion**

Menu-Driven Queries (MDQs) are a quick and efficient tool that require no programming expertise. The new functionality developed for the ADAPTABLE Supplement project provides a rapid PRO assessment to investigators. Importantly, the tool is Trial agnostic and the design allows for this functionality to be used for any future trial and any PRO information. Additionally, the temporal event functionality provides flexibility with regards to what is entered for the Index Event and the time window—any number of days before and/or after are acceptable. The structure of the PCORNet CDM provides the flexibility that has enabled us to make the tool available and applicable for future studies. Future plans for MDQs include building upon the functionality discussed and expanding the temporal event functionality to query Index Events within the Encounter table. There are also proposals to add the ability to query additional tables within the PCORNet CDM and potentially other Common Data Model schemas.

**References**