The University of Kansas Medical Center's (KUMC) Division of Medical Informatics uses i2b2 as the cornerstone for research involving clinical data for diagnosis, integrating information from separate databases to generate comprehensive patient histories. The KUMC Community Biobank and the University of Kansas Physicians, with system societies (REDCap, Biospecimen Repository), national standard format data sources (Tumor Registry – NCIACCR, Hospital Quality Indicators and Diagnoses – University HealthSystem Consortium, Social Security Death Index), are preparing our data for federated queries with other institutions: RxNorm/NDF-RT is our first focus beyond ICD9 and CPT (medications, diagnoses, and procedures respectively). HERON currently contains 725 million facts on 1.9 million patients.

**Technical Approach**

**NCI Designation**

On July 12, 2012, KU was awarded NCI Cancer Center Designation, the highest priority strategic objective for the university. Informatics contributed by enhancing i2b2 for cancer research: conducting real-time analysis via a new plugin and security layer (rgate).

**I2B2 Performance Challenges**

- As the number of observation facts increased (now over 700 million), average query time increased.
- Moving to FusionIO storage memory platform (August 2011) improved performance, but after upgrading to i2b2 1.6, query times increased significantly.
- Oracle was creating sub-optimal query plans for i2b2-generated SQL.
- After considerable analysis, tuning table statistics has provided the greatest return.
- As a result of this ongoing effort, average query times have been trending downwards.

**Motivation**

- Interoperability/Ease of Use – EPIC EMR system at KUMC has local formulary and hierarchy.
- The EPIC Hierarchy is at the clinical drug (pill size) level; not ideal for researchers. For example, instead of picking a concept for Atenolol (oral), the user must select many pill-level concepts to get all oral forms (ATENOLOL 50 MG PO TAB).
- Many drugs in the EMR don’t have therapeutic or pharmaceutical class defined and therefore are not selectable in the i2b2 front-end.

**Medications**

- Medications are organized based on VA Classes at the top-level.
- Ontology further sub-divided based on RxNorm Semantic Clinical Close Form (SCDF).
- Medications linked to RxNorm based on GCN Sequence number (First Databank Inc.) or NDC codes (National Drug Code).

**On-Going Efforts**

- 88% of medication facts are mapped to the VA Class / RxNorm hierarchy, but not all medication records in EMR have a GCN Sequence Number or an NDC code.
- Next step: augment mapping with NLP via MedEx

**Engagement and Review**

- Data Review Oversight Committee (DROC), granted Limited Data Set (government access) for all users with system access; opening the door for statistical integration within HERON.
- DROC, which includes members of medical school, WHO are invited to review the process.
- REDCap-based review process. Usage stats show number of searches performed by user and total searches performed.
- DROC members audit searches resulting in a small patient set. Audit report links to the REDCap review report.
- Ad hoc reports such at the Frontiers Participant Data Richness review.

**References and Acknowledgements**

This project is supported in part by NIH grant UI1TR000001


**Follow our progress and download code from our website at:**

http://www.fmlm.kumc.edu/Website/REDCap