**Technical Approach**

**I2B2 Performance Challenges**
- As the number of observation facts increased (now over 700 million), average query time increased.
- Motivials which maintain clinical systems (University of Kansas Hospital and the University of Kansas Physicians), university systems (REDCap, Biospecimen Repository), national standard format data sources (Tumor Registry – NAACCR, Hospital Quality Indicators and Diagnoses – University HealthSystem Consortium, Social Security Death Index). We are preparing our data for federated queries with other institutions; RxNorm/NCI-Taxonomy (IT) is our first focus beyond ICD-9 and CPT (medications, diagnoses, procedures respectively). HERO currently contains 725 millions facts on 1.9 million patients.

**Medications**

**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 Product, the user must select many pill-level concepts to get all oral forms (ATENOLOL 50 MG OD TAB).
- Many drugs in the EMR don't have therapeutic or pharmaceutical class defined and therefore are not selectable in the i2b2 front-end.

**Methodology**
- Medications are organized based on VA Classes at the top-level.
- Ontology further sub-divided based on RxNorm Semantic Clinical Dose 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 EPIC EMR have a GCN Sequence Number or an NDC code.
- Next step: augment mapping with NLP via MedEx

**Research Using HERON**

- 91 students and staff sponsored (54 sponsorship requests reviewed).
- 102 users (56 faculty, 46 staff/student).
- 6096 queries executed (3077 by Medical Informatics) as of July 13, 2012.
- 22 Data Use Requests reviewed by DROC (13 Identified data requests and 9 De-Identified data). Examples include:
  - Incidence of upper extremity DVT in ICU.
  - Double-Blind, Randomized, Historical Control Study of the Safety and Efficacy of Estrogen in Acetate Monotherapy in Subjects with Partial Epilepsy not well controlled by Current Akepileptic Medications
  - Nutrition Males Thesias: Metabolic Syndrome, Vitamin D Status and Risk of Prostate Cancer
  - Physical Therapy PD proposal: Outcomes following spine surgeries
  - Health of the Poor: ICF is an interactive, four week course offered to 4th-year medical students who define a study, collect data and present in the area of quality improvement. Example: identifying homelessness in the Emergency Department.
- 17 Frontiers Participant Registry patient data requests fulfilled for prospective trial recruitment.

**Acknowledgements**

This project is supported in part by NIH grant LIR086069.


[Link to additional information or website]

**Revised 2011 December 8**