Understanding Diagnosis Assignment from Billing Systems Relative to Electronic Health Records for Clinical Research Cohort Identification

Russ Waitman
Kelly Gerard
Daniel W. Connolly
Gregory A. Ator

Division of Medical Informatics
Department of Internal Medicine
Center for Health Informatics
Department of Otolaryngology
University of Kansas Medical Center
Kansas City, Kansas

This project is supported in part by NIH grant UL1TR000001
• The presenter has no financial relationships with commercial interests.
Integrated data repositories: electronic health records (EHR), billing systems, and quality/research registries.

- “hidden challenge”: making informed choices from this richer picture of longitudinal health.

Chart review not scalable for all studies

Work towards automated comparisons between billing, EHR and other integrated data sources to support:

- Researchers defining cohorts for retrospective analysis, study feasibility, or trial recruitment.
- Understanding how reliably different features of the EHR are used to document diagnoses.
- Motivate methods for understanding the disease progression and how multiple sources of diagnosis – during, before, and after the encounter – may provide increased support for defining the research cohort.
Acute Myocardial Infarction: A “Simple” Test Case

Incorporate Clinical, Administrative, Research Datasources

- **Inpatient and outpatient electronic medical records (Epic)**
  - Ambulatory visit details, Ordersets & Decision Support used, enhance notes extract (*pending*)
- **Professional Services Billing and Scheduling (GE IDX)**
  - Provider services (*pending*)
- **KUCC Biospecimen Shared Resource Samples Database**
  - eSample enhance (*pending*)
- **Hospital (KUH) Tumor Registry (NAACR format)**
  - enhance ontology (e.g., ER positive) (*pending*)
- **Social Security Death Master File (NIST format)**
- **Technical Charges from hospital and clinics (UHC validated format)**
  - Costs, clinics, ED (*pending*)
- **Research Data Capture (REDCap)**
  - Replicate security model
- **Clinical Research Information System (Velos)**
  - Clinical Trials Participation (*pending*)

**HERON’s current contents**

- Demographics (master patient index)
- Race/Ethnicity
- Laboratory Results
- Nursing observations/vital signs
- Clinical Diagnoses (ICD9)
- Medications (dispensed, ordered, home meds, administered)
- Physician Orders
- Procedure charges (CPT)
- Outpatient Billing diagnoses (ICD9)
- Inpatient visit/provider service
- Family/Social/Surgical Past Medical History
- Microbiology
- Allergies/Reactions
- Specimen collected
- Tumor Staging and Grade
- Diagnosis and Treatment
- Survival and Progression
- Death per Social Security Administration
- MSDRG, APDRG, LOS, Readmissions
- Technical Charge Diagnoses ICD9
- Service line, AHRQ quality and JCAHO core measures
- Triple Negative Breast Cancer Registry initial pilot completed

Status as of July 18, 2013

[https://bmi-work.kumc.edu/work/wiki/HeronProjectTimeline#July2013Planning](https://bmi-work.kumc.edu/work/wiki/HeronProjectTimeline#July2013Planning)
- contains current plan for next several monthly releases
Second Motivation: Repurposing i2b2 Infrastructure for Inpatient Quality Improvement

- i2b2 "largely" ambulatory or population/genomics focused
- i2b2 version 1.6 with same financial encounter and modifiers now useful for inpatient research and precise attribution?

- Hospital Organizational Improvement key partners
- University HealthSystem Consortium billing diagnoses often the standard for their analyses and benchmarking
KUMC’s Ontologies in i2b2
Diagnoses

Billing (IDX, UHC)
Clinical (O2 EMR)
Problem status

CERTAIN CONDITIONS ORIGINATING IN THE PERINATAL PERIOD [428,935 facts; 13,868 patients]

COMPLICATIONS OF PREGNANCY, CHILDBIRTH, AND THE PUERPERIUM [598,768 facts; 30,72

CONGENITAL ANOMALIES [412,230 facts; 42,806 patients]

DISEASES OF THE BLOOD AND BLOOD-FORMING ORGANS [1,002,319 facts; 82,424 patients]

DISEASES OF THE CIRCULATORY SYSTEM [6,284,912 facts; 201,498 patients]

Billing (IDX, UHC)
Clinical (O2 EMR)
Problem status

ACUTE RHEUMATIC FEVER [5,628 facts; 660 patients]

CEREBROVASCULAR DISEASE [484,242 facts; 33,009 patients]

CHRONIC RHEUMATIC HEART DISEASE [25,280 facts; 3,487 patients]

DISEASES OF ARTERIES, ARTERIOLES, AND CAPILLARIES [421,767 facts; 29,260 patients]

DISEASES OF PULMONARY CIRCULATION [211,836 facts; 10,998 patients]

DISEASES OF VEINS AND LYMPHATICS, AND OTHER DISEASES OF CIRCULATORY SYSTEM

HYPERTENSIVE DISEASE [2,092,069 facts; 134,783 patients]

ISCHEMIC HEART DISEASE [939,456 facts; 44,719 patients]

Billing (IDX, UHC)
Clinical (O2 EMR)

Problem status

Acute myocardial infarction [82,623 facts; 8,694 patients]

Billing (IDX, UHC)
Clinical (O2 EMR)
Problem status

Acute myocardial infarction, of anterolateral wall [399 facts; 86 patients]

Acute myocardial infarction, of inferolateral wall [151 facts; 39 patients]

Acute myocardial infarction, of inferoposterior wall [219 facts; 35 patients]

Acute myocardial infarction, of other anterior wall [1,802 facts; 324 patients]
Acute Myocardial Infarction

”Modifiers”
IMO “Synonyms” for AMI

Greater Plains Collaborative (GPC)

PCORNet goal: convert to SNOMED
Method: Cohort Characterization

- Acute myocardial infarction (AMI) ICD9 codes category 410 recorded since January 1, 2009 in the de-identified data repository compared across:
  - EHR (Epic) observations: problem list, encounter diagnoses, past medical history
  - the University HealthSystem Consortium Clinical Database containing quality measures and hospital billing codes:
    - abstracted ICD9 all diagnoses and principal diagnosis for admission
    - abstracted AHRQ quality measure for In-hospital Mortality AMI
  - ambulatory and professional charges billing system (GE IDX) used by university physician clinics.
  - the broader class of ischemic heart disease in the EHR and billing (410-414.99)
  - cardiac marker (Tropoinin I; central lab or Point of Care) result: > 0.05ng/mL
Method: analysis via i2b2 user interface

- Ran a series of i2b2 queries and recorded the counts as they varied
  - By “modifier” attribution of the diagnosis source system or other observations
  - By if the diagnoses occurred ever during the patient’s data repository record versus the “same financial encounter feature”
    - Caveat: the GE IDX clinic and professional charges don’t link at the encounter level to the Epic EMR
- Agreement is reported when during the same encounter as well as independently across the patients’ integrated records.
- Also conducted analysis for Congestive Heart Failure
University HealthSystem Consortium Clinical Database Folder from KUMC
### Acute Myocardial Infarction Diagnoses Assignment

<table>
<thead>
<tr>
<th>Diagnosis Description</th>
<th>Hospital Billing Diagnosis 1,367</th>
<th>Primary Billing Diagnosis 920</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Independent</td>
<td>Same Encounter</td>
</tr>
<tr>
<td>Clinical EHR Diagnosis, any source, 5,520</td>
<td>1,136 (83%)</td>
<td>1,058 (77%)</td>
</tr>
<tr>
<td>Encounter diagnosis 2,111</td>
<td>1027 (75%)</td>
<td>976 (71%)</td>
</tr>
<tr>
<td>Hospital problem 55</td>
<td>30 (2%)</td>
<td>22 (2%)</td>
</tr>
<tr>
<td>Medical History dx 4092</td>
<td>415 (30%)</td>
<td>0</td>
</tr>
<tr>
<td>Primary diagnosis 296</td>
<td>135 (10%)</td>
<td>42 (3%)</td>
</tr>
<tr>
<td>Principal problem 16</td>
<td>9 (1%)</td>
<td>9 (1%)</td>
</tr>
<tr>
<td>Problem List 1938</td>
<td>1,011 (74%)</td>
<td>731 (53%)</td>
</tr>
<tr>
<td>Ischemic Heart Disease (ICD9 to 410-414.99) 29,991</td>
<td>1,294 (95%)</td>
<td>1,248 (91%)</td>
</tr>
<tr>
<td>High Troponin I, or Point of Care &gt; 0.05ng/mL 8,169</td>
<td>1,302 (95%)</td>
<td>1,287 (94%)</td>
</tr>
<tr>
<td>Clinic Billing Diagnosis (IDX) 825</td>
<td>573 (42%)</td>
<td>na</td>
</tr>
<tr>
<td>AHRQ Quality In-hospital mortality AMI 910</td>
<td>910 (67%)</td>
<td>910 (67%)</td>
</tr>
</tbody>
</table>

Assignment by hospital billing records (columns) relative to other sources of diagnoses code assignment (percentages relative to hospital diagnosis).
### Congestive Heart Failure (ICD-9 428.0 to 428.9)

<table>
<thead>
<tr>
<th>Source</th>
<th>Hospital Billing Diagnosis (n = 6156)</th>
<th>Primary Billing Diagnosis (n = 1146)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Independent¹</td>
<td>Same Encounter²</td>
</tr>
<tr>
<td>Clinical Diagnosis, EHR any source (n = 9,078)</td>
<td>4,515 (73.3)</td>
<td>3,194 (51.9)</td>
</tr>
<tr>
<td>Encounter diagnosis (n = 6,913)</td>
<td>3,725 (61.5)</td>
<td>2,911 (47.3)</td>
</tr>
<tr>
<td>Hospital problem (n = 388)</td>
<td>233 (3.8)</td>
<td>87 (1.4)</td>
</tr>
<tr>
<td>Medical history diagnosis (n = 5,103)</td>
<td>2,838 (46.1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Primary diagnosis (n = 2,347)</td>
<td>1,440 (23.4)</td>
<td>134 (2.2)</td>
</tr>
<tr>
<td>Principal problem (n = 38)</td>
<td>15 (0.2)</td>
<td>5 (0.1)</td>
</tr>
<tr>
<td>Problem List (n = 4,688)</td>
<td>3,085 (50.1)</td>
<td>2,034 (33.0)</td>
</tr>
<tr>
<td>Clinic Billing Diagnosis (IDX) (n = 4,373)</td>
<td>2,884 (46.8)</td>
<td></td>
</tr>
<tr>
<td>AHRQ In-hospital HF mortality (n = 1,181)</td>
<td>1,179 (19.2)</td>
<td>1,179 (19.2)</td>
</tr>
</tbody>
</table>

- Aligned with intuition that agreement is lower for CHF during the encounter than AMI
• EMR features to identify hospital and principal problems are “unreliably” recorded.
  – Be wary of an EMR “attribute” without analysis of utilization and workflow.
  – If it doesn’t drive workflow or billing it’s often unused.

• The problem list was updated only 57% of the time for hospitalizations with AMI as the primary diagnosis (44% for CHF).
  – As expected, AMI wasn’t recorded during the encounter as past medical history but was in past medical history approximately one third of the time.
  • Due to follow up by other practices?
Future, Extended Work

- Diagnosis assignment before and after encounter
  - How are acute diagnoses are subsequently reported as past medical history?
    - Supporting continuity of care. Ideally across health systems
- Provide feedback to clinical leadership concerned with clinical documentation quality and timeliness
  - Potentially integrate “real-time” coding activities with clinical activities to enhance discharge planning
- Incorporate attribution modifiers as a component of Greater Plains Collaborative PCORNet effort for reproducible science
Diagnosis Documentation Process in Epic Post-Revenue Cycle for Patient Admitted from Emergency Department

**Hospital Encounter**

**Emergency Department**
- Patient Arrives
  - Chief Complaint Clinician PAT_ENC_RSN_VISIT
- Tests Ordered
  - Associate Diagnosis/Symptom ED MD PAT_ENC_DX
- Dx Defined/Confirmed
  - Problem List Updated
  - Admitting MD PROBLEM_LIST

**Hospital Inpatient Unit**
- Admission
  - Notes (structured free text)
  - Update Problem List MD PAT_ENC_HOSP_PROB

**Ambulatory Setting**
- Discharging MD
  - Diagnosis association for ordering outpatient test PAT_ENC_DX
  - Reconcile Problem List PAT_ENC_HOSP_PROB

**Health Information Management**
- Account Coding Complete
  - System performs coding validation to ensure account is billed w/valid DX codes

**Organizational Improvement**
- Capture Diagnosis to Sent to UHC for Benchmarking

**UHC Clinical Summary**
- (2 page overview of encounter)

<br>

**Continuity of Care Document Sent to HIE**

< 2 days

Systems compiles a list of diagnoses entered during admission when coding begins.

"Ideally feed back to MD"
Questions?

Hypothesis #1: Admin + Clinical -> Better Knowledge?

Pre-encounter factors: typically not in electronic form accessible to provider

Environment
- "toxic jungle worker"

Social Behavioral
- "exercises by jogging 2/week"

Genetic
- "poor warfarin metabolizer"

Diet, drugs
- "carrots, alcohol, BetaBlocker"

Hypothesis #2: Computer + Clinical Process -> Better Health?

Pre-encounter factors: typically not in electronic form accessible to provider

Provider

Environment
- "toxic jungle worker"

Social Behavioral
- "exercises by jogging 2/week"

Genetic
- "poor warfarin metabolizer"

Diet, drugs
- "carrots, alcohol, BetaBlocker"

"Health" or physiologic reserve

Illness/complaint

Assaults health

Bad enough to make appointment

Healthcare Encounter
- "Black box" 

"Health" or physiologic reserve

Restores health

Body’s control system + time

Vitals, assessments, diagnosis, pre-encounter factors (history), diagnosis, death

Outcomes (not typically measured)

"Health" or physiologic reserve

Observations about CDS effectiveness

Observations about EMR use

Clinical Decision Support & Information Synthesis Tools

Decisions to do something

Actions (not recorded in EMR)

Outcomes (directly measured by health care delivery system)

Collect more data: Order Tests

Ex: Laboratory, Radiology, Pathology/ exploratory procedures

Medications (orders, prescriptions)

Procedures (surgery, PT/OT, amputation)

Unregulated changes (diet)

Administrative billing activities collect "signals" as charges in parallel with clinical activities

Charges gleaned from EMR and paper

Billing system (technical charges)

Billing system (professional charges)