Clinical Phenotyping from EHRs: Opportunities and Challenges

Rachel Richesson, PhD
Duke University School of Nursing

Presented at:
NIH Symposium: Linking Disease Model Phenotypes to Human Conditions
Rockville, MD
September 10, 2015
Opportunities

• Clinician evaluation of disease & status
  – Diagnoses
  – Problems
  – Clinical notes

• Treatments, procedures, medications

• Labs

• Patient controlled data
  – Patient portals / Patient reported outcomes
  – Biometric uploads
The Office of the National Coordinator for Health Information Technology (ONC)

The Office of the National Coordinator for Health Information Technology (ONC) is at the forefront of the administration’s health IT efforts and is a resource to the entire health system to support the adoption of health information technology and the promotion of nationwide health information exchange to improve health care. ONC is organizationally located within the Office of the Secretary for the U.S. Department of Health and Human Services (HHS).

ONC is the principal Federal entity charged with coordination of nationwide efforts to implement and use the most advanced health information technology and the electronic exchange of health information. The position of National Coordinator was created in 2004, through an Executive Order, and legislatively mandated in the Health Information Technology for Economic and Clinical Health Act (HITECH Act) of 2009.

ONC’s mission includes:

- Promoting development of a nationwide Health IT infrastructure that allows for electronic use and exchange of information that:
  - Ensures secure and protected patient health information
  - Improves health care quality
  - Reduces health care costs
  - Informs medical decisions at the time/place of care
  - Includes meaningful public input in infrastructure development
  - Improves coordination of care and information among hospitals, labs, physicians, etc.
  - Improves public health activities and facilitates early identification/rapid response to public health emergencies
  - Facilitates health and clinical research
  - Promotes early detection, prevention, and management of chronic diseases
  - Promotes a more effective marketplace
  - Improves efforts to reduce health disparities
- Providing leadership in the development, recognition, and implementation of standards and the certification of Health IT products;
- Health IT policy coordination;
- Strategic planning for Health IT adoption and health information exchange; and
- Establishing governance for the Nationwide Health Information Network.

Federal Register Notice

August 13, 2010: Statement of Organization, Functions, and Delegations of Authority: Office of the National Coordinator for Health and Information Technology Correction [PDF - 40 KB]

December 1, 2009: Organization, Functions, and Delegations of Authority: Office of the National Coordinator for Health Information Technology [PDF - 49 KB]

Celebrating the First Anniversary of the HITECH Act and Looking to the Future: February 2010 [PDF - 852 KB]
Meaningful Use as a Building Block

Stage 1 MU

- Basic EHR functionality, structured data
- Privacy & security protections

Stage 2 MU

- Care coordination
- Patient informed
- Structured data utilized
- Privacy & security protections

PCMHs 3-Part Aim

- Data utilized to improve delivery and outcomes
- Evidenced based medicine
- Registries for disease management
- Privacy & security protections

Stage 3 MU

ACOs

- Improved population health
- Enhanced access and continuity
- Patient self management
- Patient engaged, community resources
- Patient centered care coordination
- Team based care, case management
- Registries to manage patient populations
- Privacy & security protections

Use information to transform

Improve access to information

Utilize technology to gather information
Hospitals Demonstrating Meaningful Use

91-95%

Office-Based Providers Demonstrating Meaningful Use

54%

Patient Portal Use in Large Health Plan

Figure 1. Age Not a Major Driver of Portal Adoption
Proportion of Patients with Portal Accounts
By Age Group, Among Patients with Encounters in the Past 30 Days

Source: athenaResearch.
Sample: 973k patients visiting practices offering the athenaCommunicator portal.

Outstanding Challenges

• EHRs designed to support clinical care
• Completeness & accuracy vary
Outstanding Challenges

• Still not standardized
• 100+ EHR vendor products
• Coding systems used differently
• ICD-10 transition
• Researchers do not control EHR design or documentation/coding practices
Clinical Phenotype Definitions

- Specifications for **identifying patients or populations with a given characteristic or condition of interest from EHRs** using data that are routinely collected in EHRs or ancillary data sources.
- Can support research (cohort selection; study screening) and personalized medicine
- Include widely adopted coding systems
  - ICD-9-CM
  - CPT
  - SNOMED CT
  - LOINC
  - RxNorm
  - NDC
Example

Diabetes defined as:1

- one inpatient discharge diagnosis (ICD-9-CM 250.x, 357.2, 366.41, 362.01-362.07)

or any combination of two of the following events occurring within 24 months of each other:

- A1C \geq 6.5\% (48 \text{ mmol/mol})
- fasting plasma glucose \geq 126 \text{ mg/dl} (7.0 \text{ mmol/L})
- random plasma glucose \geq 200 \text{ mg/dl} (11.1 \text{ mmol/L})
- 2-h 75-g OGTT \geq 200 \text{ mg/dl}
- outpatient diagnosis code (same codes as inpatient)
- anti-hyperglycemic medication dispense (see details below)
- NDC in associated list
- ...etc., etc...

The eMERGE Network

The mapping of the human genome has enabled new exploration of how genetic variations contribute to health and disease. To better realize this promise, researchers must now determine ways in which genetic make-up gives some individuals a greater chance of becoming sick with chronic conditions such as diabetes, Alzheimer’s, or heart disease. The goal of gaining this knowledge is to translate it to bedside practice and ultimately improve patient care.

The Electronic Medical Records and Genomics (eMERGE) Network is a national consortium organized by NHGRI to develop, disseminate, and apply approaches to research. It combines DNA biorepositories with electronic medical record (EMR) systems for large-scale, high-throughput genetic research with the ultimate goal of returning genomic testing results to patients in a clinical care setting. The Network is currently exploring more than a dozen phenotypes (with 13 additional electronic algorithms having already been published). Various models of returning clinical results have been implemented or planned for pilot at sites across the Network. Themes of bioinformatics, genomic medicine, privacy and community engagement are of particular relevance to eMERGE.

What makes eMERGE unique?

Each center participating in the Network is studying the relationship between genome-wide genetic variation and a common human trait. Such studies commonly involve testing hundreds of thousands of genetic variants called single nucleotide polymorphisms (SNPs) throughout the genome in people with and without the trait. A number of such studies are reporting an association between disease and a person’s genetic make-up, but those studies are typically costly and take a long time to complete.

The eMERGE model is exploring use of data from the EMR – clinical systems that represent an alternative methodology. Electronic medical records are one of the most exciting potential sources of data to better understand the genetic basis of disease. Each center site has EMR data linked to genetic samples obtained in the course of existing clinical care from residual tissue or blood samples. In the eMERGE model, there is no need to actively collect study population. Cases and controls are quickly and consistently identified from the EMR and are readily available. This approach is both cost-effective and time-efficient. More detailed information on the phenotypes being explored in eMERGE can be found on the PheKB and other freely downloadable Resources page.

In addition, eMERGE focuses on ethical, legal, social, and policy issues such as privacy and...
<table>
<thead>
<tr>
<th>Title</th>
<th>Groups</th>
<th>Institutions</th>
<th>Data and Methods</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atrial Fibrillation - Demonstration</td>
<td>Vanderbilt - SD/RD Group</td>
<td>Vanderbilt University</td>
<td>CPT Codes, ICD 9 Codes, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Project</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac Conduction (QRS)</td>
<td>eMERGE Phenotype WG</td>
<td>Vanderbilt University</td>
<td>CPT Codes, ICD 9 Codes, Laboratories, Medications, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Cataracts</td>
<td>eMERGE Phenotype WG</td>
<td>Marshfield Clinic Research</td>
<td>CPT Codes, ICD 9 Codes, Medications, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Foundation</td>
<td></td>
<td>Foundation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clopidogrel Poor Metabolizers</td>
<td>Denny’s Group at Vandy, VESPA - Vanderbilt, Electronic Systems for Pharmacogenomic Assessment</td>
<td></td>
<td>CPT Codes, ICD 9 Codes, Laboratories, Medications, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Crohn's Disease - Demonstration Project</td>
<td>Vanderbilt - SD/RD Group</td>
<td>Vanderbilt University</td>
<td>ICD 9 Codes, Medications, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Dementia</td>
<td>eMERGE Phenotype WG</td>
<td>Group Health Cooperative</td>
<td>ICD 9 Codes, Medications</td>
<td>Final</td>
</tr>
<tr>
<td>Diabetic Retinopathy</td>
<td>eMERGE Phenotype WG</td>
<td>Marshfield Clinic Research</td>
<td>CPT Codes, ICD 9 Codes, Medications, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Foundation</td>
<td></td>
<td>Foundation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug Induced Liver Injury</td>
<td>eMERGE Phenotype WG</td>
<td>Columbia University</td>
<td>ICD 9 Codes, Laboratories, Medications, Natural Language Processing</td>
<td>Final</td>
</tr>
<tr>
<td>Rheumatoid Arthritis - Demonstration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Other Sources for Clinical Phenotypes

• Clinical Classifications Software, “AHRQ Bundles”
• CMS Chronic Conditions Warehouse
• Quality Net (CMS and Joint Commission)
• Mini-Sentinel
• OMOP/OHDSI
• SHARPn
• .....
A comparison of phenotype definitions for diabetes mellitus

Rachel L Richesson,1 Shelley A Rusincovitch,2 Douglas Wixted,3 Bryan C Batch,4 Mark N Feinglos,5 Marie Lynn Miranda,5 W Ed Hammond,2,6 Robert M Califf,3,7 Susan E Spratt4

ABSTRACT

Objective This study compares the yield and characteristics of diabetes cohorts identified using heterogeneous phenotype definitions. Furthermore, standard phenotype definitions can streamline the development of patient registries from healthcare data, and enable consistent inclusion criteria to support regional surveillance and prevention initiatives.

Table 1 Data domain criteria used in selected phenotype definitions

<table>
<thead>
<tr>
<th>Phenotype definitions:</th>
<th>Data domain criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICD-9-CM 250.xx</td>
<td></td>
</tr>
<tr>
<td>CMS CCW</td>
<td></td>
</tr>
<tr>
<td>NYC A1c Registry</td>
<td></td>
</tr>
<tr>
<td>Diabetes-associated medications</td>
<td></td>
</tr>
<tr>
<td>DDC</td>
<td></td>
</tr>
<tr>
<td>SUPREME-DM</td>
<td></td>
</tr>
<tr>
<td>eMERGE†</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ICD-9-CM 250.x0 and 250.x2 (excludes type 1 specific codes)</th>
<th>Expanded ICD-9-CM Codes (249.xx, 357.2, 362.0x, 366.41)</th>
<th>HbA1c</th>
<th>Fasting glucose</th>
<th>Random glucose</th>
<th>Abnormal OGTT</th>
<th>Diabetes-associated medications*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICD-9-CM 250.xx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMS CCW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYC A1c Registry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes-associated medications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUPREME-DM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eMERGE†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Medications vary by phenotype definition and are listed for each in the supplementary appendix (available online only).
†The eMERGE phenotype definition consists of five case scenarios with varying combinations of criteria. Any instance of type 1 specific codes (ie, 250.x1, 250.x3) results in the exclusion of the patient.

- = Sole criteria.
△ = Optional criteria, one of many.
☆ = Distinction made between inpatient and outpatient context.
\ = Distinction made for multiple instances and/or time points.

CMS CCW, Centers for Medicare and Medicaid Services Chronic Condition Data Warehouse; DDC, Durham Diabetes Coalition; eMERGE, electronic medical records and genomics; HbA1c, hemoglobin A1c; ICD-9-CM, International Classification of Disease, revision 9, clinical modification; NYC, New York City; OGTT, oral glucose tolerance test; SUPREME-DM, Surveillance, Prevention, and Management of Diabetes Mellitus.
Clinical Trial Reporting

Multiple phenotype definitions:

**SUPREME-DM Phenotype**

**Definition:**
- Adult Durham Population patients who meet **ONE OR MORE** of the following criteria during a **DukeMed** encounter between *2007-2011*:
  - One or more instances of the specified ICD-9-CM diagnosis codes (see table 7) on an **inpatient** encounter
  - OR 2 or more instances of the specified ICD-9-CM diagnosis codes (see table 7) on **outpatient** encounters on separate days
  - OR 1 or more instances of active stand-alone medication (see table 8) reported during outpatient medication reconciliation
  - OR 1 or more Oral Glucose Tolerance Test (OGTT) 2-hour 75g result >= 200 mg/dl where there is NO DIAGNOSIS CODE on the same encounter indicating pregnancy (V22, V23)*
  - OR 2 or more hemoglobin A1c results >= 6.5% on **2 different days** within 730 day span
  - OR 2 or more fasting glucose results >= 126 mg/dl on **2 different days** within 730 day span
  - OR 2 or more random glucose results >= 200 mg on **2 different days** within 730 day span
  - OR within a 750 day span on **2 different days**:
    - Fasting glucose results >= 126 mg/dl
    - OR Random glucose results >= 200 mg
  - OR within a 750 day span (can be same day):
    - Hemoglobin A1c results >= 6.5%
    - OR Random glucose results >= 200 mg

**Abnormal Lab Results**

**Source:** Laboratory results

**Definition:**
- Adult Durham Population patients who meet **ONE OR MORE** of the following criteria during a **DukeMed** encounter between *2007-2011*:
  - One or more instances of hemoglobin A1c results >= 6.5%
  - OR one or more fasting glucose results >= 126 mg/dl within 365 day span
  - OR one or more random glucose results >= 200 mg/dl within 365 day span.

**Abnormal HbA1c (NCY A1c Registry Definition)**

**Source:** Glucose and hemoglobin laboratory results

**Definition:**
- Adult Durham Population patients who meet **ONE OR MORE** of the following criteria during a **DukeMed** encounter between *2007-2011*:
  - One or more instances of hemoglobin A1c results >= 6.5%
Multicenter Study Comparing Case Definitions Used to Identify Patients with Chronic Obstructive Pulmonary Disease

Valentin Prieto-Centurion¹, Andrew J. Rolle¹, David H. Au², Shannon S. Carson³, Ashley G. Henderson³, Todd A. Lee⁴, Peter K. Lindauer⁵,6, Mary A. McBurnie⁷, Richard A. Mularski⁷, Edward T. Naureckas⁸, William M. Vollmer⁷, Binoy J. Joese⁹, and Jerry A. Krishnan¹,⁹; on behalf of the CONCERT Consortium

¹Division of Pulmonary, Critical Care, Sleep and Allergy and ⁴Department of Pharmacy Systems, Outcomes and Policy, University of Illinois at Chicago, Chicago, Illinois; ²University of Washington/VA Puget Sound, Seattle, Washington; ³Division of Pulmonary and Critical Care Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; ⁵Department of Medicine and Center for Quality of Care Research, Baystate Medical Center, Springfield, Massachusetts; ⁶Tufts University School of Medicine, Boston, Massachusetts; ⁷The Center for Health Research, Kaiser Permanente, Portland, Oregon; ⁸Section of Pulmonary and Critical Care, University of Chicago Medicine, Chicago, Illinois; and ⁹Population Health Sciences Program, University of Illinois Hospital and Health Sciences System, Chicago, Illinois

Am J Respir Crit Care Med. 2014 Nov 1;190(9):989-95. doi: 10.1164/rccm.201406-1166OC.
Table 2. Clinical Characteristics of Patients Who Met and Did Not Meet the Clinical Trial Reference Standard

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Sample (n = 998)</th>
<th>Clinical Trial Reference Standard</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Yes* (n = 560)</td>
<td>No† (n = 438)</td>
</tr>
<tr>
<td>Comorbid conditions, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>76</td>
<td>74</td>
<td>78</td>
</tr>
<tr>
<td>Hypertension</td>
<td>66</td>
<td>63</td>
<td>69</td>
</tr>
<tr>
<td>Heart failure</td>
<td>18</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>23</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>19</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Stroke</td>
<td>15</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Depression</td>
<td>42</td>
<td>36</td>
<td>50</td>
</tr>
<tr>
<td>Arthritis</td>
<td>36</td>
<td>33</td>
<td>41</td>
</tr>
<tr>
<td>Diabetes</td>
<td>28</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td>Cancer history</td>
<td>23</td>
<td>26</td>
<td>19</td>
</tr>
<tr>
<td>Anemia</td>
<td>28</td>
<td>26</td>
<td>30</td>
</tr>
<tr>
<td>Kidney disease</td>
<td>20</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Dementia</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Dyspnea at rest (Borg), %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0, no dyspnea</td>
<td>52</td>
<td>54</td>
<td>50</td>
</tr>
<tr>
<td>0.5–2, slight</td>
<td>38</td>
<td>38</td>
<td>37</td>
</tr>
<tr>
<td>≥3, moderate to very severe</td>
<td>10</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Spirometry, post-bronchodilator, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV₁/FVC &lt;70%</td>
<td>61</td>
<td>100</td>
<td>11</td>
</tr>
<tr>
<td>FEV₁ &lt;80% predicted</td>
<td>72</td>
<td>86</td>
<td>55</td>
</tr>
<tr>
<td>6-minute-walk distance, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance walked &lt;350 m</td>
<td>53</td>
<td>52</td>
<td>54</td>
</tr>
</tbody>
</table>

Patients who met the trial reference standard are more likely to have airflow obstruction by spirometry but report being less dyspneic. Patients who met the reference standard also have different prevalence of comorbidities. For example, they are more likely to have hypertension, heart failure, and depression. Data for 6-minute-walk distance missing in 9% patients (9% and 10%) and dyspnea scores missing in 8% patients (8% and 9%) in those who met and did not meet the clinical trial reference standard, respectively.

*(A + D + E + G) and †(B + C + F) in Figure 2.
Table 3. Characteristics Associated with Meeting the Clinical Trial Reference Standard

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race (vs. white)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.37 (0.26–0.53)*</td>
</tr>
<tr>
<td>Other</td>
<td>0.52 (0.27–1.00)</td>
</tr>
<tr>
<td>Education (vs. high school or less)</td>
<td></td>
</tr>
<tr>
<td>College/professional degree</td>
<td>0.38 (0.26–0.56)*</td>
</tr>
<tr>
<td>Some college</td>
<td>0.68 (1.06–2.03)*</td>
</tr>
<tr>
<td>BMI, kg/m² (vs. normal)</td>
<td></td>
</tr>
<tr>
<td>&lt;18.5 (underweight)</td>
<td>4.00 (1.27–12.50)*</td>
</tr>
<tr>
<td>25–29.99 (overweight)</td>
<td>0.87 (0.58–1.30)</td>
</tr>
<tr>
<td>≥30 (obese)</td>
<td>0.51 (0.35–0.75)*</td>
</tr>
<tr>
<td>Depression (yes vs. no)</td>
<td>0.53 (0.40–0.71)*</td>
</tr>
<tr>
<td>Diabetes (yes vs. no)</td>
<td>0.67 (0.48–0.93)*</td>
</tr>
<tr>
<td>Cancer (yes vs. no)</td>
<td>1.47 (1.05–2.08)*</td>
</tr>
</tbody>
</table>

Definition of abbreviations: BMI = body mass index; CI = confidence interval.

Clinical trial reference standard \((A + D + E + G)\) versus others \((B + C + F)\) in Figure 2. Multivariable logistic regression model that included characteristics listed in Tables 1 and 2 (characteristics significantly associated with meeting the trial reference standard). Results indicate that patients who are black (vs. white), with college or higher (vs. high school or less) education, obese (vs. normal weight), with depression, or diabetes are less likely to meet the trial reference standard. Patients with a history of cancer and underweight patients (vs. normal weight) are more likely to meet the trial reference standard. Hosmer-Lemeshow goodness-of-fit test \((P \text{ value } = 0.17)\) demonstrates adequate model fit.

\(^*P < 0.05.\)
Upcoming Events
Grand Rounds March 7: Bray Patrick-Lake (CTTI; PCORnet Executive Committee member), Sue Sheridan (PCORI), and Sean Tunes (CMTIP). Patient Engagement in Infrastructure Development
Secretary’s Advisory Committee for Human Research Protections (SACHRP): March 12-13

Grand Rounds March 14: TBD

Subscribe to our mailing list: nih-collaboratory@dm.duke.edu.

Knowledge Repository
View Collaboratory products, resources, publication references, etc.

Collaboratory News
Stop CRC featured on NPR Health Blog
02/26/2014: Gloria Coronado, PhD, was recently featured on the NPR Health Blog discussing the Stop CRC study.

Joe Selby writes perspective piece for the New England Journal of Medicine on PCORI
02/13/14: Joe Selby, MD, MPH, Executive Director of PCORI, published a perspective piece in the latest issue of the New England Journal of Medicine on lessons learned in PCORI’s 3-year history.

First patient enrolled in Collaboratory trial
01/13/14: The TIME Demonstration Project, led the University of Pennsylvania’s Laura Dember, MD, has enrolled its first patient.

Featured Topics
Articles, presentations, and other products related to specific topics of interest.
- Regulatory Update related to SUPPORT Trial
- Demonstration Projects - Regulatory and Ethics Discussions
- NIH Collaboratory Communication Channels Chart

Educational Presentations: Archives
02-21-14: Sharon Terry
Participant Engagement: Tools to Meet People Where They Are

02-14-14: Eric Larson
Engaging Health Systems in Research Partnerships

PubMed Related Articles
<table>
<thead>
<tr>
<th>URL</th>
<th>Publication Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rescuing clinical trials in the United States and beyond: A call for action.</td>
<td>2013/06</td>
<td>To promote consensus around the solutions needed to address the adverse trends in clinical research, the Duke Clinical Research Institute convened stakeholders from academia, industry, and government. This article summarizes the proceedings.</td>
</tr>
<tr>
<td>Rapid, responsive, relevant (R3) research: a call for rapid learning health research enterprise</td>
<td>2013/05</td>
<td>To produce more rapid, responsive, and relevant research, we propose approaches that increase relevance via greater stakeholder involvement, speed research via innovative designs, streamline review processes.</td>
</tr>
<tr>
<td>Human subjects protections in community-engaged research: a research ethics framework</td>
<td>2010/03</td>
<td>This new framework for exploring the risks in community-engaged research can help academic researchers and community partners ensure the mutual respect that community-engaged research requires.</td>
</tr>
</tbody>
</table>
PCORnet: The National Patient-Centered Clinical Research Network

The Patient-Centered Outcomes Research Institute (PCORI) is supporting the development of PCORnet, the National Patient-Centered Clinical Research Network, to create a large, highly representative, national network for conducting clinical outcomes research.

PCORnet will transform clinical research by engaging patients, care providers, and health systems in collaborative partnerships to improve healthcare and advance medical knowledge. By bringing research and patient care together, this innovative health data network will be able to explore the questions that matter most to patients and their families. Read more …

Patient-focused drug development moves into the mainstream

Posted March 16, 2015, Applied Clinical Trials

Building on the experience gained from FDA meetings soliciting patient perspectives, sponsors, and stakeholders, the patient-focused drug development process has moved into the mainstream. Read more …
## Coverage and Precision of Rare Disease Names (n=5,333)

<table>
<thead>
<tr>
<th></th>
<th>Equivalent match</th>
<th>Broader match (found)</th>
<th>Broader match (projected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICD-9-CM</td>
<td>697 (13%)</td>
<td>2055 (39%)</td>
<td>2569 (48%)</td>
</tr>
<tr>
<td>ICD-10-CM</td>
<td>1386 (26%)</td>
<td>919 (17%)</td>
<td>1635 (31%)</td>
</tr>
<tr>
<td>SNOMED CT</td>
<td>2848 (53%)</td>
<td>Not estimated</td>
<td>Not estimated</td>
</tr>
</tbody>
</table>

“Grouper” codes that contain multiple rare disease concepts (less precise)

<table>
<thead>
<tr>
<th># rare diseases included in code</th>
<th># ICD-9-CM codes</th>
<th># ICD-10-CM codes</th>
<th># SNOMED CT codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1081 (62%)</td>
<td>1403 (73%)</td>
<td>3311 (85%)</td>
</tr>
<tr>
<td>2</td>
<td>319</td>
<td>328</td>
<td>478</td>
</tr>
<tr>
<td>3</td>
<td>125</td>
<td>88</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>68</td>
<td>45</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>43</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>&gt; 5</td>
<td>117</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td># codes including &gt; 1 disease (%)</td>
<td>672 (38%)</td>
<td>526 (27%)</td>
<td>598 (15%)</td>
</tr>
</tbody>
</table>

Examples
- 208 rare diseases included under 759.89 Other specified congenital anomalies
- 22 rare diseases included under Q82.8 Other specified congenital malformations of skin
- 5 rare diseases included under 28835009 Retinitis pigmentosa

SNOMED CT

• Most comprehensive, multilingual clinical terminology in the world
• Used in > 50 countries
• Meaningful Use requires use of SNOMED CT in the EHR for problem lists, procedures.
• SNOMED CT is better suited for clinical data capture because:
  – Better content coverage
  – Clinically oriented
  – Flexible data entry and retrieval
Use Cases and Coding Systems for Rare Diseases

HPO and ORDO for “deep phenotyping” of undiagnosed disorders in specialty or genetics clinics.

Link to OMIM and GO and for molecular diagnosis.

MeSH for linkage to the biomedical literature and clinical practice guidelines (e.g., InfoButton, CDSS).

Linkage to patient-directed health information (e.g., Medline Plus search with MeSH synonyms).

Electronic Health Record Systems

Encode with SNOMED CT for documenting diagnoses or “problems”

Reimbursement
ICD-9-CM, ICD-10-CM

Public Health Surveillance
ICD-10

Quality Measurement
SNOMED CT

Interventional Research
SNOMED CT, MedDRA; plus new data collection using PhenX and LOINC

Query SNOMED CT for networked research networks and observational research.

Link to OMIM and GO and for molecular diagnosis.
## Multiple Terminologies

<table>
<thead>
<tr>
<th>1. Outpatient COPD encounter</th>
<th>ICD-9 (or corresponding SNOMED) codings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>491.xx - chronic bronchitis</td>
</tr>
<tr>
<td></td>
<td>492.xx - emphysema</td>
</tr>
<tr>
<td></td>
<td>493.2x - chronic obstructive asthma</td>
</tr>
<tr>
<td></td>
<td>496.xx - chronic airway obstruction</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Inpatient COPD encounter</th>
<th>ICD-9 (or corresponding SNOMED) codings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>490.xx - bronchitis, not specified as acute or chronic</td>
</tr>
<tr>
<td></td>
<td>491.xx - chronic bronchitis</td>
</tr>
<tr>
<td></td>
<td>492.xx - emphysema</td>
</tr>
<tr>
<td></td>
<td>493.2x - chronic obstructive asthma</td>
</tr>
<tr>
<td></td>
<td>494.xx - bronchiectasis</td>
</tr>
<tr>
<td></td>
<td>495.xx - extrinsic allergic alveolitis</td>
</tr>
<tr>
<td></td>
<td>496.xx - chronic airway obstruction</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Hospitalized respiratory failure (primary) with secondary COPD coding</th>
<th>518.81 acute respiratory failure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>518.82 other pulmonary insufficiency not elsewhere classified</td>
</tr>
<tr>
<td></td>
<td>518.84 acute and chronic respiratory failure</td>
</tr>
</tbody>
</table>

| 4. Problem List | Coded fields ICD-9 / variant “V” coding or SNOMED – COPD, emphysema, chronic bronchitis, chronic airway obstruction |

| 5. PFT demonstrating likely obstructive lung disease | Evidence of pre or post-bronchodilator obstruction defined by an FEV1 / FVC ratio < 0.7 |

<table>
<thead>
<tr>
<th>6. Documented COPD-related drug dispensing, order or medication listing</th>
<th>Beta-agonists by nebulized delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta-agonists, Inhalers - Long acting</td>
</tr>
<tr>
<td></td>
<td>Ipratropium &amp; Tiotropium</td>
</tr>
<tr>
<td></td>
<td>Inhaled Corticosteroid (ICS)</td>
</tr>
<tr>
<td></td>
<td>Combivent (Ipratropium+Albuterol)</td>
</tr>
<tr>
<td></td>
<td>Combination Inhaled Beta-agonists + ICS (ie Advair)</td>
</tr>
</tbody>
</table>

**Medication classes**

- Beta-agonists by nebulized delivery
- Beta-agonists, Inhalers - Long acting
- Ipratropium & Tiotropium
- Inhaled Corticosteroid (ICS)
- Combivent (Ipratropium+Albuterol)
- Combination Inhaled Beta-agonists + ICS (ie Advair)

**Need to link to**

- RxNorm
- NDC
Figure 1. The various subdomains integrated in the UMLS.

- Clinical repositories
- Genetic knowledge bases
- OMIM
- MeSH
- Biomedical literature
- SNOMED
- NCBI Taxonomy
- GO
- UWDA
- Genome annotations
- Anatomy
- Model organisms
- Other subdomains
- Human Phenotype Ontology

Olivier Bodenreider Nucl. Acids Res. 2004;32:D267-D270
Possible actions...

• Link specialized terminologies with UMLS

• Recognize importance of clinical interface terminology (SNOMED CT or HPO)
  – Advocate for standards (requirements or incentives)
  – Promote use of SNOMED CT in healthcare systems

• Expand coverage of diseases in SNOMED CT

• Laboratory data critical

• Coordinate and communicate across disciplines!
Acknowledgments

• This work was made possible by the National Institutes of Health (NIH) Common Fund, through a cooperative agreement (U54 AT007748) from the Office of Strategic Coordination within the Office of the NIH Director, and the Patient Centered Outcomes Research Institute.

• The views presented here are solely the responsibility of the author and do not necessarily represent the official views of the National Institutes of Health or PCORI.

• Kin Wah Fung and Olivier Bodenreider (National Library of Medicine) supported exploration of rare disease names in UMLS.