PSQ Core - Update

Rachel Richesson
W. Ed Hammond
Meredith Nahm Zozus

April 20, 2015

PSQ Core Focus Areas

- **<u>P</u>henotypes**
  - Support and discussion for individual groups
  - General guidance (Living Text, web page, journals and conferences)
  - Grow a library or collection of phenotype definitions

- **Data <u>S</u>tandards**
  - Propose "standard" phenotype definitions and approaches to EHR data
  - How to get these used and reported in studies (the "Table 1 project")

- **Data <u>Q</u>uality**
  - [white paper](#) (Meredith Zozus, lead), part of online Living Text
  - Watching PCORI Data Quality Group, and DQ Common Data Model
  - **The use of population-level data is essential to explore, measure, and report "data quality" so that the results can be appropriately interpreted.**
  - Need adequate data and methods to detect the likely and genuine variation between populations at different trial sites and/or intervention groups.

Challenge - Presenting Baseline Characteristics for Clinical Trial Reporting

Multiple possible phenotypes:

Patient characteristics:

Table 1. Patient Demographics and Baseline Characteristics

Characteristic	No. (%) of Patients[a]	
	Gentamicin-Collagen Sponge (n = 753)	Control (n = 749)
Patient demographics		
Age, median (IQR), y	64.2 (58.0-71.5)	64.9 (57.2-72.1)
White race	688 (91.4)	683 (91.2)
Weight, median (IQR), kg	98.0 (86.1-113.0)	98.8 (85.0-111.1)
Body mass index, median (IQR)	33.1 (30.2-37.2)	32.8 (30.0-36.2)
Body mass index >30	574 (76.2)	563 (75.2)
Male sex	530 (70.4)	530 (70.8)
Medical history		
History of hypertension	659 (87.5)	659 (88.0)
History of diabetes	493 (65.5)	513 (68.5)
Current or history of smoking	458 (60.8)	450 (60.1)
Current smoking	136 (29.7)	123 (27.3)
History of chronic obstructive pulmonary disease	117 (15.5)	107 (14.3)
History of peripheral vascular disease	105 (13.9)	89 (11.9)
Previous median sternotomy	52 (6.9)	42 (5.6)
History of TIA or stroke	77 (10.2)	81 (10.8)
History of myocardial infarction	233 (31.0)	245 (32.7)
History of congestive heart failure	89 (11.8)	90 (12.0)
History of hyperlipidemia	619 (82.2)	607 (81.0)
Steroid use ≤1 mo prior to surgery	28 (3.7)	33 (4.4)
Receiving dialysis preoperatively	4 (0.5)	2 (0.3)
Preoperative diagnostic values		
Left ventricular ejection fraction, median (IQR), %	55 (45-60)	55 (45-60)
Serum glucose, median (IQR), mg/dL	125 (101-160)	124 (103-167)
Serum hemoglobin A_{1c}, median (IQR), %	6.5 (5.9-7.6)	6.6 (5.9-7.7)
Hematocrit, median (IQR), %	39 (36-42)	39 (36-42)
Serum creatinine, median (IQR), mg/dL	1.0 (0.9-1.3)	1.0 (0.9-1.2)
Preoperative core temperature, median (IQR), °C	97.6 (97.0-98.2)	97.7 (97.0-98.2)
Preoperative hospital stay, median (IQR), d	1.0 (0-3.0)	1.0 (0-3.0)
Parsonnet risk score, median (IQR)[b]	9.0 (6.0-14.5)	9.0 (6.0-16.0)

Abbreviations: IQR, interquartile range; TIA, transient ischemic attack.
SI conversion factors: To convert creatinine to μmol/L, multiply by 88.4; glucose to mmol/L, multiply by 0.0555.
[a] Unless otherwise indicated.
[b] Theoretical range is 0 to 148; 50% in Parsonnet et al[11] had a score between 0 and 9.

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[3]
- 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)[4]
- 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 730 day span on 2 different days:
  - Fasting glucose results >= 126 mg/dl
  - AND Random glucose results >= 200 mg
- OR within a 730 day span (can be same day):
  - Hemoglobin A1c results >= 6.5%
  - AND Fasting glucose results >= 126 mg/dl

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:
Glycated 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%

Collaboratory Approaches

Current approach:

- Support search and evaluation of existing phenotypes

- Facilitate documentation

- Share definitions

- Share implementation information and results

- Link to other resources

- Promote as "best available"

- Work toward standards

Collaboratory Approaches

Current approach:

- Support search and evaluation of existing phenotypes

- Facilitate documentation

- Share definitions

- Share implementation information and results

- Link to other resources

- Promote as "best available"

- Work toward standards

Supporting activities:

- List of authoritative sources
- Create documentation "templates":
  - Logic
  - Supporting evidence
  - Implementation guidance

- Develop methods for scientific validation studies

- Data quality recommendations for any pragmatic trials using EHR data

Presentation for phenotypes (in progress)

Selection and planning: *Implementation:*

Recommendations on Collaboratory website

→ *link to* →

Justification and guidance for use in Pragmatic Trials

→ *link to* →

Human readable phenotype, collaboration, versioning, public dissemination

---

Phenotype Definitions Used

NIH Collaboratory
Rethinking Clinical Trials®
Health Care Systems Research Collaboratory

NIH Collaboratory   About Us ▾   Demonstration Projects ▾   Cores ▾   News ▾
The Living Textbook   Grand Rounds   News A

Cores › Phenotypes, Data Standards, and Data Quality
H Collaboratory Cores and Working Groups

Populations:
Patients w/ chronic pain
Patients w/ imaging studies for lower back pain
Patients who are candidates for CRC screening

….

Confounders or Risks:
Diabetes
Hypertension

…

Outcomes:
Mortality
Suicide attempt

NIH Collaboratory *Rethinking Clinical Trials™*
Health Care Systems Research Collaboratory

Type 2 Diabetes Mellitus Phenotype Definitions

From the NIH Collaboratory Phenotypes, Data Standards, and Data Quality Core

Available at: https://www.nihcollaboratory.org/Pages/Knowledge-Repository.aspx

Background: The Phenotypes, Data Standards, and Data Quality Core of the NIH Health Care Systems Research Collaboratory is developing a series of recommendations for the collection/query of data from electronic health records (EHRs) and/or ancillary systems for person characteristics and clinical features to support standardized reporting of baseline characteristics of research populations in interventional and observational studies.

Purpose of this document: This document represents our synthesis of existing phenotype definitions that have been used in diabetes research and population health activities. Using guidelines for the evaluation of existing phenotypes, our informatics and EHR phenotyping experience, and specialized clinical/research expertise, we suggest a suite of phenotype definitions, each appropriate for a particular purpose. The following is our recommendation, complete with a justification and supporting information and resources, for explicit EHR-derived phenotype definitions for diabetes. However, neither the Collaboratory nor the NIH has formally endorsed these definitions or their use in the data collection or reporting of this condition at this time (see disclaimer).

Audience: This document and supporting information is directed to clinical researchers and research sponsors who are making decisions about the data to use for studies. These documents should provide specifications and guidance that will assist researchers in making informed and deliberate choices about EHR data to use in research studies. The supporting

---

PheKB   a knowledgebase for discovering phenotypes from electronic medical records

Home | Phenotypes | Implementations | Resources | Groups | eMERGE Network | Contact Us

🏠 › Phenotypes

📋 Phenotypes

Group		Include Methods	Exclude Methods		Mine Only	
- Any -					- Any -	Apply

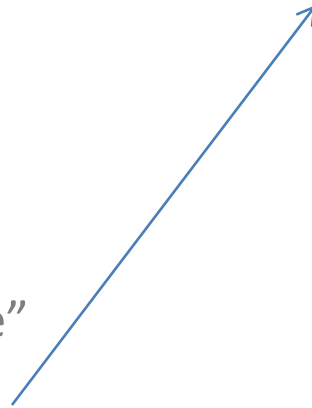| Title	Groups	Institutions	Data and Methods	Status
Atrial Fibrillation - Demonstration Project	Vanderbilt - SD/RD Group	Vanderbilt University	CPT Codes, ICD 9 Codes, Natural Language Processing	Final
Cardiac Conduction (QRS)	eMERGE Phenotype WG	Vanderbilt University	CPT Codes, ICD 9 Codes, Laboratories, Medications, Natural Language Processing	Final
Cataracts	eMERGE Phenotype WG	Marshfield Clinic Research Foundation	CPT Codes, ICD 9 Codes, Medications, Natural Language Processing	Final
Clopidogrel Poor Metabolizers	Denny's Group at Vandy, VESPA - Vanderbilt Electronic Systems for Pharmacogenomic Assessment		CPT Codes, ICD 9 Codes, Laboratories, Medications, Natural Language Processing	Final
Crohn's Disease - Demonstration Project	Vanderbilt - SD/RD Group	Vanderbilt University	ICD 9 Codes, Medications, Natural Language Processing	Final
Dementia	eMERGE Phenotype WG	Group Health Cooperative	ICD 9 Codes, Medications	Final
Diabetic Retinopathy	eMERGE Phenotype WG	Marshfield Clinic Research Foundation	CPT Codes, ICD 9 Codes, Medications, Natural Language Processing	Final

---

In the future….

link to ↓

Standard code lists (VSAC) or executable code

Collaboratory Approaches

Current approach:

- Support search and evaluation of existing phenotypes
- Facilitate documentation
- Share definitions
- Share information
- Link to other resources
- Promote as "best available"
- **Work toward standards**

- Evaluate and compare
- Facilitate use
- Explore incentives
- Advocate:
  - Research sponsors
  - SDOs
  - Policy makers

Acknowledgments

Monique Anderson	Tammy Reece
Lesley Curtis	Shelley Rusincovitch
W. Ed Hammond	Michelle Smerek
Cindy Kluchar	**Meredith Nahm Zozus**

Members of the Phenotype Core of the NIH Collaboratory:

Alan Bauck, Kaiser Permanente

Denise Cifelli, U. Penn.

John Dickerson, Kaiser Permanente Northwest

Pedro Gozalo, , Brown Univ. School of Public Health & Providence VA Health Services Research Service

Bev Green, Group Health

Chris Helker, U. Penn

Michael Kahn, Children's Hospital of Colorado

Reesa Laws, Kaiser Permanente

Melissa Leventhal, University of Colorado Denver

John Lynch, Connecticut Institute for Primary Care Innovation

Rosemary Madigan, U. Penn

Vincent Mor, Brown Univ. School of Public Health & Providence VA Health Services Research Service

George "Holt" Oliver, Parkland Health and Hospital System (UT Southwestern)

Jon Puro, OCHIN

Jerry Sheehan, National Library of Medicine

Greg Simon, Group Health

Kari Stephens, U. of Washington

Erik Van Eaton, U. of Washington