Practical Development and Implementation of EHR Phenotypes

NIH Collaboratory Grand Rounds
Friday, November 15, 2013

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

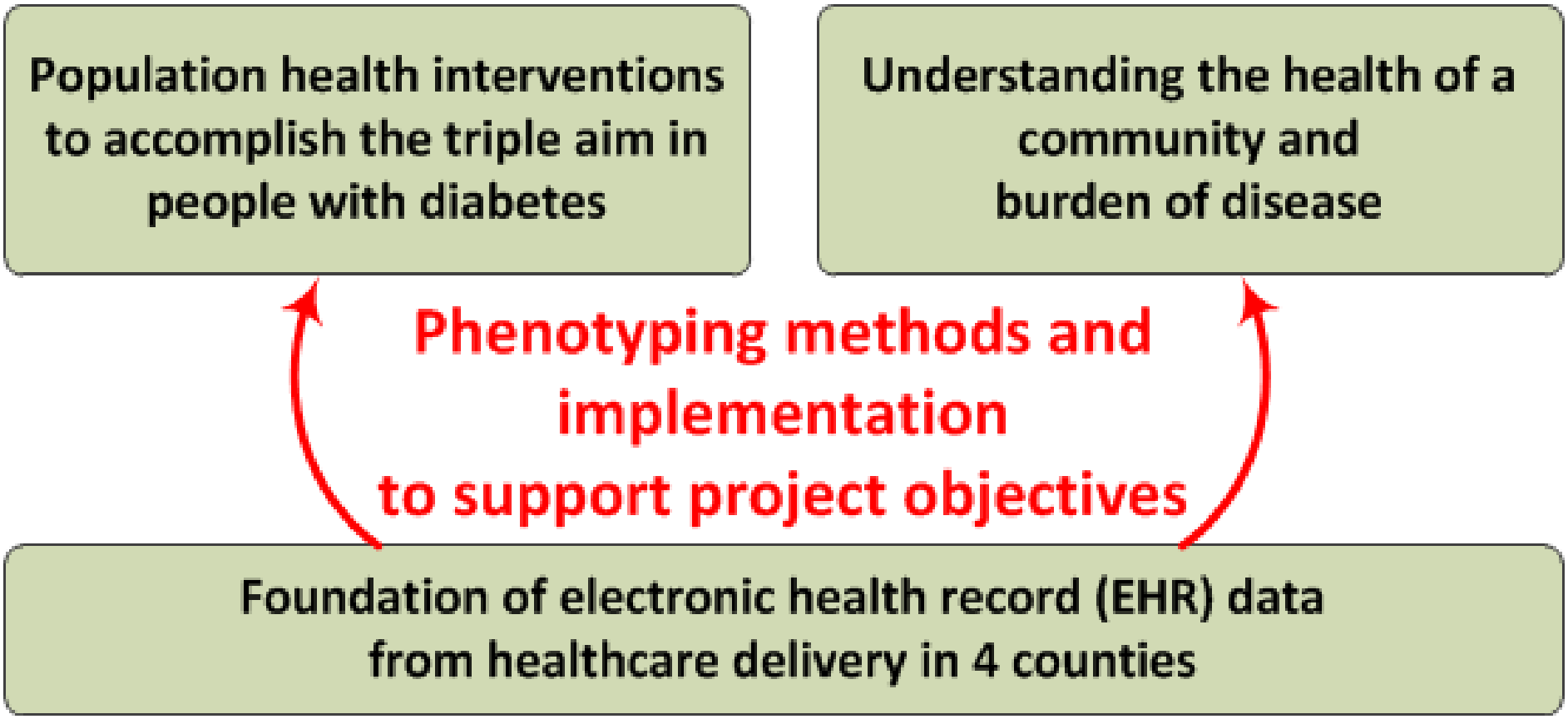
The Southeastern Diabetes Initiative (SEDI)

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Setting the Context



Risk Prediction and Intervention: The Need for Clinical Risk Factors and Outcomes

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Why Computable Phenotypes?

- Correct “disease” identification
- Several downstream implications
  - Estimation of incidence
  - Study design: include and exclude
  - Identification of “risk factors”
  - Effect estimation
  - Who (how) to treat and who to spare
  - Bias due to incorrect disease specification

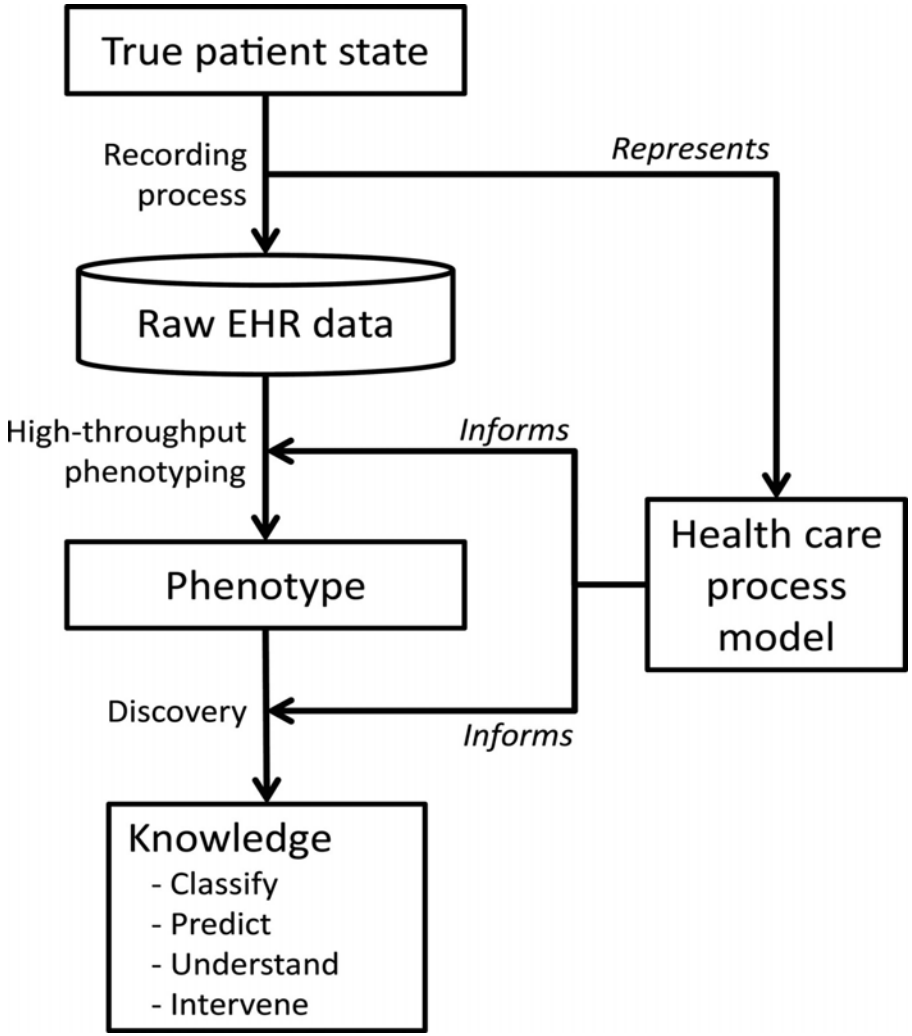
Attribution: Paramita Saha Chaudhuri, PhD

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Phenotyping and discovery.



Hripcsak G , and Albers D J J Am Med Inform Assoc 2013;20:117-121

SEDI Medical-Social Risk Algorithm Drives Intervention

- Different intensities of intervention
    - High-intensity clinical teams vs. lower-intensity community-based teams
- Different modes of intervention
    - Patient basis, neighborhood basis, community basis
- Targeted intervention
    - Stratifying patients based on risk, both at patient and neighborhood levels

The Diabetes Phenotype Comparison

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Problem Statement

- EHR-driven computable phenotypes exist and are an important source of knowledge
  - Electronic Medical Records and Genomics (eMERGE) Network and Phenotype Knowledge Base
  - Entities such as the Center for Medicare & Medicaid Services
  - Many others
- How should we recognize, document, implement, and validate authoritative source phenotypes?
- How should we evaluate the best fit and utility of phenotypes, especially applied to population health management?

Initial

Problem Statement

- Which patients in a 5-year EHR dataset have diabetes?

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

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,[4] Marie Lynn Miranda,[5] W Ed Hammond,[2,6] Robert M Califf,[3,7] Susan E Spratt[4]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/amiajnl-2013-001952).

[1]Duke University School of Nursing, Durham, North Carolina, USA
[2]Applied Informatics Research, Duke Health Technology Solutions, Durham, North Carolina, USA
[3]Duke Translational Medicine Institute, Durham, North Carolina, USA
[4]Division of Endocrinology, Metabolism and Nutrition, Department of Medicine, Duke University School of Medicine, Durham, North Carolina, USA
[5]Department of Pediatrics, School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michig
[6]Duke Inform North
[7]Divisi Depart Univer Durha

ABSTRACT

Objective This study compares the yield and characteristics of diabetes cohorts identified using heterogeneous phenotype definitions.

Materials and methods Inclusion criteria from seven diabetes phenotype definitions were translated into query algorithms and applied to a population (n=173 503) of adult patients from Duke University Health System. The numbers of patients meeting criteria for each definition and component (diagnosis, diabetes-associated medications, and laboratory results) were compared.

Results Three phenotype definitions based heavily on ICD-9-CM codes identified 9–11% of the patient population. A broad definition for the Durham Diabetes Coalition included additional criteria and identified 13%. The electronic medical records and genomics, NYC A1c Registry, and diabetes-associated medications definitions, which have restricted or no ICD-9-CM criteria, identified the smallest proportions of patients (7%). The

populations. Furthermore, standard phenotype definitions can streamline the development of patient registries from healthcare data, and enable consistent inclusion criteria to support regional surveillance and the identification of rare disease complications. An understanding of the populations generated from various phenotype definitions will inform standard methods for identifying diabetes cohorts, facilitate the rapid generation of patient registries and research datasets with uniform sampling criteria, and enable comparative and aggregate analysis. This descriptive study presents and compares the size and characteristics of patient populations retrieved using different phenotype definitions adopted from prominent diabetes registries and research networks munity intervention program in ou federal reporting standards.

BACKGROUND AND SIGNIFICANCE

Diabetes diagnosis and managemen

sease with rent etiolog nagement s M) is the in the US rcise, oral mellitus. C

Revised DDC Diabetes Phenotype

Source:
Developed by Durham Diabetes Coalition (DDC), January 2013, revised in May 2013.

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 5**) on any type of encounter (inpatient, outpatient, ED)
- OR one or more active medications associated with DM treatment reported during outpatient medication reconciliation (see **Table 6**)
- OR two or more hemoglobin A1c results >= 6.5% within 365 day span
- OR two or more fasting glucose results >= 126 mg/dl within 365 day span
- OR two or more random glucose results >= 200 mg/dl within 365 day span
- OR within the same 365-day span, at least two of the following:
  - Hemoglobin A1c result >= 6.5%
  - Fasting glucose result >= 126 mg/dl
  - Random glucose result >= 200 mg/dl
  - Oral Glucose Tolerance Test (OGTT) 2-hour 75g result >= 200 mg/dl[2]

Table 5: Revised DDC ICD-9-CM Codes Indicating Type 2 Diabetes: 249.xx, 250.xx, 357.2, 362.01-07, 366.41, but not including type 1 specific codes (250.x1 and 250.x3)

DIAGNOSIS_CODE	DIAGNOSIS_LONG_DESC
249.00	SECONDARY DIABETES MELLITUS WITHOUT MENTION OF COMPLICATION, NOT STATED AS UNCONTROLLED, OR UNSPECIFIED
249.01	SECONDARY DIABETES MELLITUS WITHOUT MENTION OF COMPLICATION, UNCONTROLLED
249.10	SECONDARY DIABETES MELLITUS WITH KETOACIDOSIS, NOT STATED AS UNCONTROLLED, OR UNSPECIFIED
249.11	SECONDARY DIABETES MELLITUS WITH KETOACIDOSIS, UNCONTROLLED

Table 1 Data domain criteria used in selected phenotype definitions

Phenotype definitions:	Data domain criteria: ICD-9-CM 250.xx	ICD-9-CM 250.x0 and 250.x2 (excludes type 1 specific codes)	Expanded ICD-9-CM Codes (249.xx, 357.2, 362.0x, 366.41)	HbA1c	Fasting glucose	Random glucose	Abnormal OGTT	Diabetes-associated medications*
ICD-9-CM 250.xx	●							
CMS CCW	▲*||		▲*||					
NYC A1c Registry				●				
Diabetes-associated medications								●
DDC		▲	▲	▲||	▲||	▲||	▲||	▲
SUPREME-DM	▲*||		▲*||	▲||	▲||	▲||	▲	▲
eMERGE†		●*||		▲	▲	▲		▲

*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.

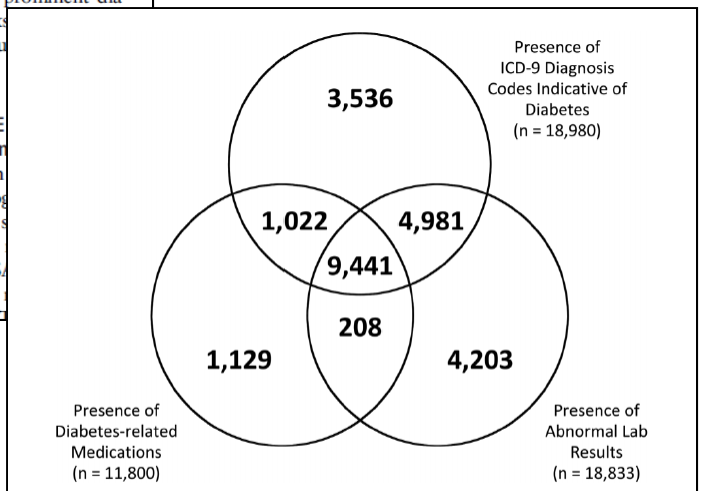


Figure 1 Overlap of diabetes cohorts identified from different categories of phenotype eligibility criteria; n=24 520 patients identified by criteria from any of the three categories.

Richesson RL, Rusincovitch SA, Wixted D, Batch BC, Feinglos MN, Miranda ML, Hammond WE, Califf RM, Spratt SE. A Comparison of Phenotype Definitions for Diabetes Mellitus. J Am Med Inf Assoc 2013 (epub ahead of print). http://www.ncbi.nlm.nih.gov/pubmed/24026307

Individual Cohort Yields

Table 2 Demographic characteristics of cohorts returned from selected diabetes phenotype definitions

Characteristic	DUHS reference population	Phenotype definition						
		ICD-9-CM 250.xx Codes	CMS CCW (full ICD-9 set)	NYC A1c Registry	Diabetes-associated medications	DDC phenotype	SUPREME-DM	eMERGE
Purpose for phenotype	–	Billing	Health services research	Care management	–	Community-wide intervention	Epidemiology; community-wide intervention	Genetic research
Type of diabetes targeted	–	All	All	All	*T2DM preferred	*T2DM preferred	All	†T2DM exclusive
Age—year‡ (mean±SD)†	41.7±17.5	56.1±15.8	56.8±15.5	56.2±15.1	54.1±15.3	55.6±16.3	56.6±15.9	57.3±15.4
Female sex: # and (%)	99 695 (57%)	10 644 (56%)	9185 (56%)	6812 (56%)	6933 (59%)	12 603 (57%)	10 681 (56%)	6524 (56%)
No of encounters§ (mean±SD)†	20±33.5	46±57.1	49±58.6	54±59.3	54±60.4	46±56.3	48±57.9	45±52.5
Length of time (in days) between first and last patient encounter (mean±SD)†	861±675.9	1252±587.6	1295±558.4	1365±524.5	1394±500.5	1224±595.9	1257±576.1	1258±579.4
Total patients identified	173 503	18 893	16 320	12 182	11 800	22 050	18 958	11 620
% Reference population identified	n/a	11%	9%	7%	7%	13%	11%	7%

*Project focus or intent is for T2DM populations, but phenotype does not aggressively eliminate T1DM patients.
†Patients with indications of T1DM are specifically excluded.
‡Age at the beginning of the observation period, 1 January 2007.
§Within observation period, 1 January 2007–31 December 2011.
CMS CCW, Centers for Medicare and Medicaid Services Chronic Condition Data Warehouse; DDC, Durham Diabetes Coalition; DUHS, Duke University Health System; eMERGE, electronic medical records and genomics; HbA1c, hemoglobin A1c; ICD-9-CM, International Classification of Disease, revision 9, clinical modification; NYC, New York City; SUPREME-DM, Surveillance, Prevention, and Management of Diabetes Mellitus; T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus.

Challenge: Representing and Comparing Phenotype Criteria

Table 1 Data domain criteria used in selected phenotype definitions

Phenotype definitions:	Data domain criteria							
	ICD-9-CM 250.xx	ICD-9-CM 250.x0 and 250.x2 (excludes type 1 specific codes)	Expanded ICD-9-CM Codes (249.xx, 357.2, 362.0x, 366.41)	HbA1c	Fasting glucose	Random glucose	Abnormal OGTT	Diabetes-associated medications*
ICD-9-CM 250.xx	●							
CMS CCW	▲ ✳ \\		▲ ✳ \\					
NYC A1c Registry				●				
Diabetes-associated medications								●
DDC		▲	▲	▲ \\	▲ \\	▲ \\	▲ \\	▲
SUPREME-DM	▲ ✳ \\		▲ ✳ \\	▲ \\	▲ \\	▲ \\	▲	▲
eMERGE†		● ✳ \\		▲	▲	▲		▲

*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.

Simple Phenotype Criteria Example: ICD-9-CM Diagnosis Category 250.xx

Source:
ICD-9-CM diagnosis code 250 with any degree of specificity in the fourth and fifth decimal precision (250.xx).

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 ICD-9-CM diagnosis code 250.xx (see **table 1**) for any type of encounter (inpatient, outpatient, ED)

Complex Phenotype Criteria Example: 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
- OR within a 730 day span (can be same day):
  - Hemoglobin A1c results >= 6.5%
  - AND Random glucose results >= 200 mg

Very Complex Phenotype Criteria Example: eMERGE (NW) Phenotype

Definition:
Adult Durham Population patients who meet **ONE OR MORE** of the following criteria during a DukeMed encounter between 2007-2011:

- Case 1:
  - Zero instances of T1 diagnosis codes (see **Table 9**) on any type of encounter (inpatient, outpatient, ED)
  - AND 1 or more instances of T2 diagnosis codes (see **Table 10**) on any type of encounter (inpatient, outpatient, ED)
  - AND has BOTH T1 meds and T2 meds (see **Table 11** and **Table 12**)
  - AND t2 med date is PRIOR TO t1 med date
- OR Case 2:
  - Zero instances of T1 diagnosis codes (see **Table 9**) on any type of encounter (inpatient, outpatient, ED)
  - AND 1 or more instances of T2 diagnosis codes (see **Table 10**) on any type of encounter (inpatient, outpatient, ED)
  - AND has NO instances of T1 meds (see **Table 11)**
  - AND has one or more T2 meds (see **Table 12**)
- OR Case 3:
  - Zero instances of T1 diagnosis codes (see **Table 9**) on any type of encounter (inpatient, outpatient, ED)
  - AND 1 or more instances of T2 diagnosis codes (see **Table 10**) on any type of encounter (inpatient, outpatient, ED)
  - AND has NO instances of T1 meds (see **Table 11)**
  - AND has NO or more T2 meds (see **Table 12**)
  - AND has at least one abnormal lab:
    - Hemoglobin A1c result >= 6.5%
    - Fasting glucose result >= 125 mg/dl
    - Random glucose result >= 200 mg/dl
- OR Case 4:
  - Zero instances of T1 diagnosis codes (see **Table 9**) on any type of encounter (inpatient, outpatient, ED)
  - AND 0 instances of T2 diagnosis codes (see **Table 10**) on any type of encounter (inpatient, outpatient, ED)
  - AND has 1 or more instances of T2 meds (see **Table 12**)
  - AND has at least one abnormal lab:
    - Hemoglobin A1c result >= 6.5%
    - OR Fasting glucose result >= 125 mg/dl
    - OR Random glucose result >= 200 mg/dl
- OR Case 5:
  - Zero instances of T1 diagnosis codes (see **Table 9**) on any type of encounter (inpatient, outpatient, ED)
  - AND 2 or more instances of T2 diagnosis codes made on at least TWO SEPARATE dates on any type of encounter (inpatient, outpatient, ED)
  - AND has 1 or more instances of T1 meds (see **Table 11)**
  - AND has 0 instances of T2 meds (see **Table 12**)

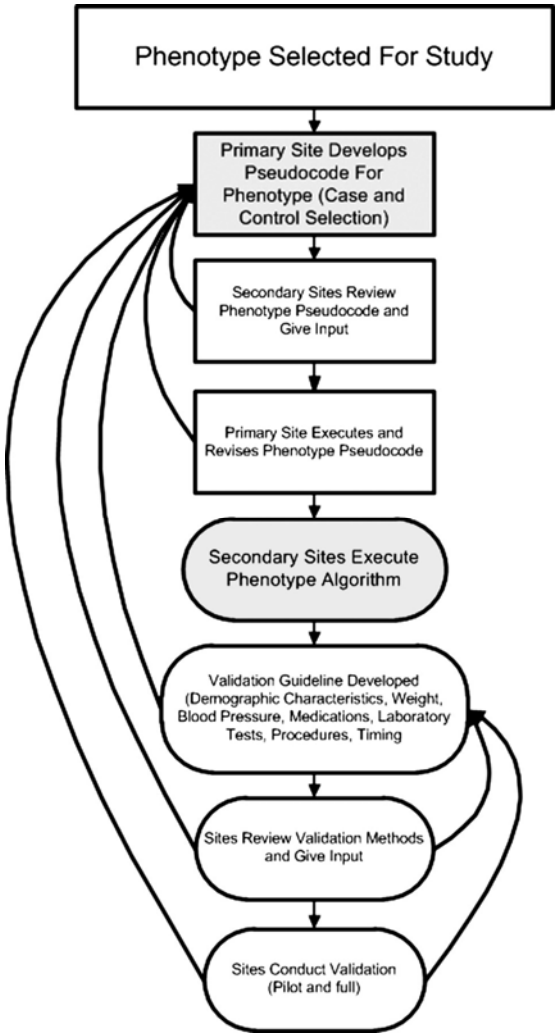
Phenotypes Development: A Pragmatic Approach

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

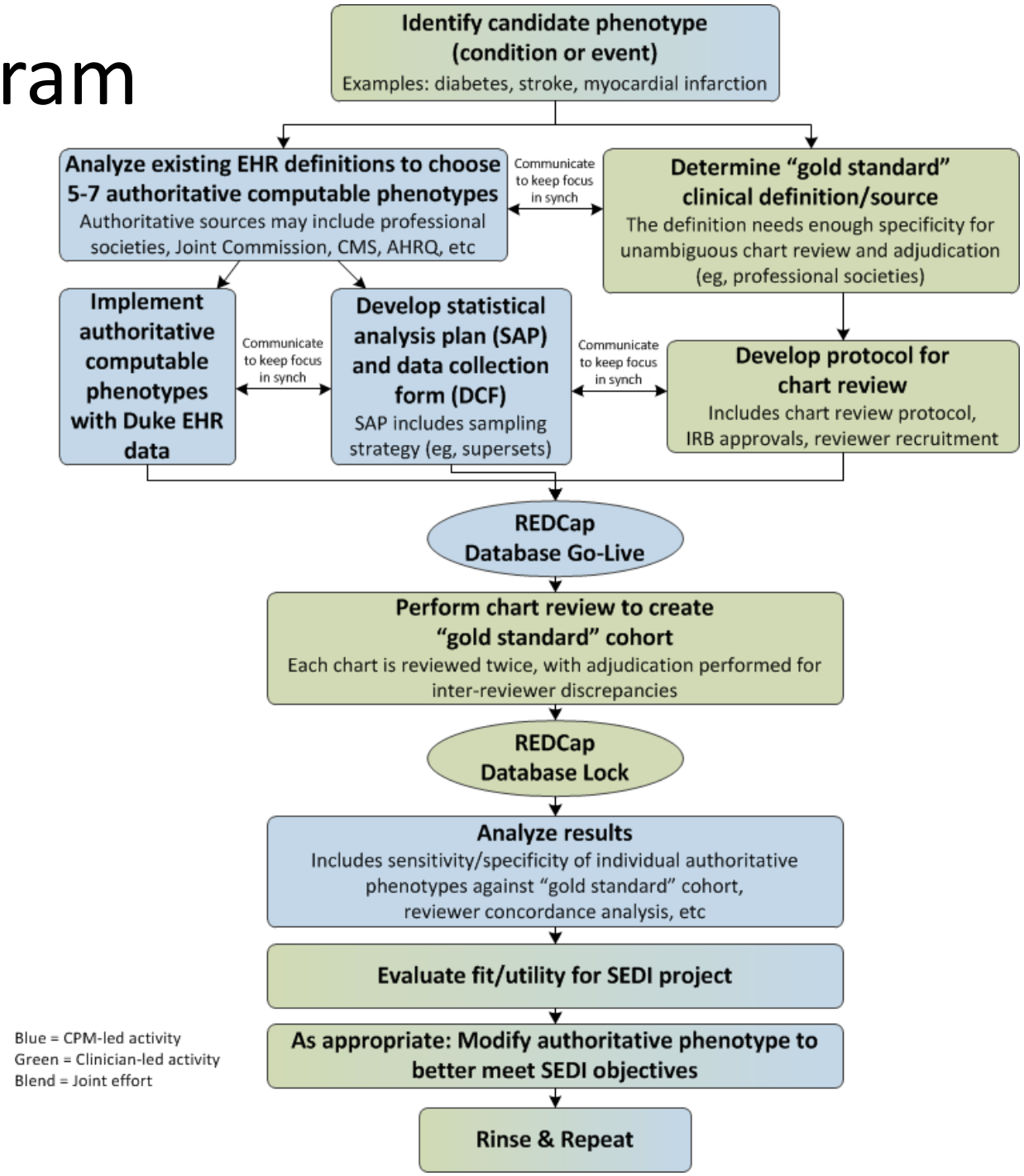
Center for Predictive Medicine
Duke Clinical Research Institute

Phenotype development and validation.



Newton K M et al. J Am Med Inform Assoc 2013;20:e147-e154

Process Diagram



Attribution: Center for Predictive Medicine

Recognizing Authoritative Sources

Table 1: Primary Phenotype Sources

Source	*Comments*
Clinical Classifications Software (CCS), also known as AHRQ Bundles	Only based upon diagnosis codes, but very large listing of conditions; this is the basis for most early SEDI variables. http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp
CMS Chronic Conditions Warehouse (CCW)	Only based upon diagnosis codes and procedure codes; clinical review to date has felt that inclusion logic can be overly broad. https://www.ccwdata.org/web/guest/condition-categories http://www.ncbi.nlm.nih.gov/pubmed/21649659
Mini-Sentinel	Exhaustively researched definitions, but limited number of phenotypes represented. http://www.mini-sentinel.org/assessments/diagnoses_and_medical_procedures/default.aspx
eMERGE Network and PheKB phenotypes library	Probably the most well-recognized phenotyping source at present, but limited number of phenotypes represented; should be carefully evaluated because core mission of genomic studies can result in exclusionary logic inappropriate for the SEDI population health focus. http://www.phekb.org/phenotypes http://www.ncbi.nlm.nih.gov/pubmed/21269473
Quality Net (joint effort of CMS and Joint Commission)	Separates measures between inpatient basis and outpatient basis. Go to the "specifications manual" option; the appendixes contain specific listings of ICD-9 code tables, medication tables, and CPT codes. This is one of the only CPT code groupings that we've seen so far (CPT licensure is very restrictive), but QualityNet only includes for outpatient context. https://www.qualitynet.org/
National Drug File Reference Terminology (NDF-RT)	Search on a term (eg, diabetes) using "contains" and "name/code" specifiers. The results tab for "view all" contains the "may_treat" relationship of conditions to drugs. http://nciterms.nci.nih.gov/ncitbrowser/pages/vocabulary.jsf?dictionary=National%20Drug%20File%20-%20Reference%20Terminology
Professional society guidelines	These are an important source for definitions of abnormal laboratory results and specific ranges, which are often not represented in other definitions. Examples: American Diabetes Association, National Kidney Foundation, American College of Cardiology
Major and well-recognized clinical trials and registries using EHR data to identify cohorts	Clinical and expert guidance can be important for identification of these pivotal trials; another potential technique might be to limit results to high-impact journals via a PubMed search.

Table 2: Secondary Phenotype Sources

Source	*Comments*
Joint Commission	**The CMS/Joint Summit QualityNet is generally the better source, not using the Joint Commission directly.** This organization evaluates hospital adherence with federal regulations, and publishes a specifications manual for inpatient quality measures. Appendix A.1 lists the definitions for specific conditions, mostly based upon ICD-9. A limitation is that these definitions are centered on inpatient admissions, and may not be applicable in an outpatient setting. http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx
World Health Organization (WHO) Global Burden of Disease	**In general, this may be useful for mental health, but probably not helpful for most clinical condition phenotypes.** The Global Burden of Disease classifications include both ICD-9 and ICD-10 diagnosis code groupings. See "cause-specific documentation" for individual conditions (eg, cerebrovascular disease, diabetes mellitus, etc). The diagnosis codes are not granular (eg, it just lists 250 for diabetes mellitus), due to global application, and the clinical conditions are very broad. May be somewhat out of date; it appears that the classifications date back to 2000; the last formal GBD update appears to have been 2004, although this is difficult to ascertain from their website. However, there are a lot of mental health classifications, which may be useful. http://www.who.int/healthinfo/global_burden_disease/data_sources_methods/en/index.html
Meaningful Use	*This area needs further research. Does MU publish specific phenotypes for disease conditions? Most documentation appears related to attestation of technical capacities, especially in stage 1, not clinical definitions.* http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Meaningful_Use.html

Attribution: Center for Predictive Medicine

Evaluating Existing Definitions

Phenotype Overview: Acute Myocardial Infarction (research by Maria V. Grau-Sepulveda)

Clinical Definition Source: Joint ESC/ACCF/AHA/WHF Task Force for the Universal Definition of MI[1]

Table 1: Authoritative Phenotype Comparison

Source	*Evaluation of Prevalence vs. Incidence*	*EHR Data Subject Areas*	*Phenotype Comments*	*Phenotype Encounter Basis*

AHRQ Bundles (Clinical Classifications Software)[2]	Prevalence	ICD-9 Diagnoses	• Broad definition
			• AMI diagnosis codes:
			  o initial episode
			  o subsequent episode
			  o unspecified episode
				Any encounter

CMS Chronic Conditions Warehouse[3]	Incidence	ICD-9 Diagnoses Encounter Basis	• Only AMI initial episode codes	Inpatient basis, first/second diagnosis code
Mini-Sentinel #1 (AMI/Anti-Diabetic Agents)[4]	Incidence	ICD-9 Diagnoses Encounter Basis Death Data	• AMI initial/unspecified episode codes	Inpatient basis, first diagnosis code Also includes death w/i one day of ED visit with ischemic disease codes
Mini-Sentinel #2 (Validation of AMI Cases)[5]	Incidence	ICD-9 Diagnoses Encounter Basis	• AMI initial/unspecified episode codes • Does not include death criteria	Inpatient basis, first diagnosis code
CMS/Joint Summits QualityNet (Yale models for AMI and HF)[6] Joint Commission identification of AMI[7]	Incidence	ICD-9 Diagnoses Encounter Basis	• AMI initial/unspecified episode codes	Inpatient basis, first diagnosis code

[1] http://www.escardio.org/guidelines-surveys/esc-guidelines/GuidelinesDocuments/Guidelines_Univ_Def_Myocardial_Infarc_FT.pdf
[2] http://www.hcup-us.ahrq.gov/toolssoftware/ccs/AppendixASingleDX.txt
[3] https://www.ccwdata.org/cs/groups/public/documents/document/ccw_conditionreferencelist2011.pdf
[4] http://www.mini-sentinel.org/work_products/Assessments/Mini-Sentinel_AMI-and-Anti-Diabetic-Agents_Protocol.pdf
[5] http://mini-sentinel.org/work_products/Validation_HealthOutcomes/Mini-Sentinel-Validation-of-AMI-Cases.pdf
[6] https://www.qualitynet.org/dcs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1228889871496&blobheader=multipart%2Foctet-stream&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3Bfilename%3D2.1+AMI_4.2a.pdf&blobcol=urldata&blobtable=MungoBlobs
[7] http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx

Attribution: Maria Grau-Sepulveda, MD, MPH

SEDI Core Data Domains

1. Patient Demographics
2. Encounters
3. Diagnoses
4. Procedures
5. Lab Results
6. Vital Signs
7. Medications
8. Social History

Kahl M, Dunston FG, Morris LM, Rusincovitch SA. Traceability in Healthcare Data Sharing Projects Through the Use of Data Warehousing Artifacts: Methods from the Southeastern Diabetes Initiative (SEDI). HDWA (Healthcare Data Warehouse Association) 2013 Annual Conference. October 1-3, 2013, Scottsdale, Arizona. Abstract: Poster presentation.

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Healthcare Workflows and EHR Data

Data Reflective of Biomedical Phenomena:	Data Reflective of Diagnostic Processes:
• Laboratory result values • Vital sign measures • Direct physiological measures (such as EKG, pulmonary function tests, etc) • Pathology specimens • Images	• Diagnosis codes (includes professional billing, technical billing, medical coding) • Problem lists • Clinical narrative related to diagnosis (including pathology and imaging reports)
Data Reflective of Behavior, Functioning, or Experience of Symptoms:	**Data Reflective of Treatment Decisions:**
• Patient-reported outcomes • Social and family history • Other instruments addressed to patient	• Provider orders (including medications) • Procedure codes • Procedure reports (such as surgery reports) • Clinical narrative relating to treatment plans

Patient-centered context, but mediated by provider decisions of diagnostic testing and exposure to health system

Healthcare-centric context, but mediated by billing processes, medical coding conventions, and healthcare EHR system platform

A work in progress with Greg Simon, Michelle Smerek, and Rachel Richesson

Contact Information

Shelley A. Rusincovitch, Senior Informatics Analyst
Health Intelligence and Research Services
Duke Health Technology Solutions (DHTS)

Phone: 919-668-5954 / shelley.rusincovitch@duke.edu

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Supplemental Slides

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Selected SEDI References

Califf RM, Sanderson I, Miranda ML. The future of cardiovascular clinical research: informatics, clinical investigators, and community engagement. JAMA 2012 7;308(17):1747-8.
http://www.ncbi.nlm.nih.gov/pubmed/23117773

Miranda ML, Ferranti J, Strauss B, Neelon B, Califf RM. Geographic health information systems: a platform to support the 'triple aim'. Health Aff (Millwood) 2013;32(9):1608-15.
http://www.ncbi.nlm.nih.gov/pubmed/24019366

Granger BB, Rusincovitch SA, Avery S, Batch BC, Dunham AA, Feinglos MN, Kelly K, Pierre-Louis M, Spratt SE, Califf RM. Missing signposts on the roadmap to quality: a call to improve medication adherence indicators in data collection for population research. Frontier Pharmacol 2013;4(13):1-13.
http://www.ncbi.nlm.nih.gov/pubmed/24223556

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute

Center for Predictive Medicine

Paramita Saha Chaudhuri, PhD

Maria Grau-Sepulveda, MD, MPH

Eric Laber, PhD

Nick Meyer

Ben Neely, MS

Charlotte Nelson, MS

Michael Pencina, PhD

Shelley Rusincovitch

DukeMedicine

SOUTHEASTERN DIABETES INITIATIVE

Center for Predictive Medicine
Duke Clinical Research Institute