# Factors and Approaches to Mapping Laboratory Results in PCORnet

**Michelle M. Smerek, Elisa Priest, Dr.PH, S. Trent Rosenbloom, MD, MPH, FACMI,**

**Jon E. Puro, MPA, Pedro Rivera, MSCS, Shelley A. Rusincovitch,**

**Rahul Jain, MPH, CPHIMS, and Keith Marsolo, PhD**

*AMIA 2016 Joint Summits on Translational Science*
*March 21-24, 2016, San Francisco, California*

pcornet

AMIA

# Disclosure

- Neither I nor my life partner have any relevant financial relationships with commercial interests.

# Learning Objectives

After participating in this activity, the learner should be better able to:

- Articulate factors that make mapping source data system lab results to the PCORnet Common Data Model "common measures" challenging
- Articulate approaches/strategies for addressing these challenges

pcornet

AMIA

# PCORnet Overview

- A “network of networks”
- Composed of:
  - 13 Clinical Data Research Network
  - 20 Patient-Powered Research Networks
- Distributed network where the questions go to the data

PCORnet seeks to improve the nation’s capacity to conduct clinical research by creating a large, highly representative, national patient-centered network that supports more efficient clinical trials and observational studies.

pcornet

AMIA

# PCORnet Common Data Model v3.0

| DEMOGRAPHIC |
|---|
| **PATID** |
| BIRTH_DATE |
| BIRTH_TIME |
| SEX |
| HISPANIC |
| RACE |
| BIOBANK_FLAG |

**Fundamental basis**

| ENROLLMENT |
|---|
| **PATID** |
| **ENR_START_DATE** |
| ENR_END_DATE |
| CHART |
| **ENR_BASIS** |

| DISPENSING |
|---|
| **DISPENSINGID** |
| **PATID** |
| PRESCRIBINGID (optional) |
| **DISPENSE_DATE** |
| **NDC** |
| DISPENSE_SUP |
| DISPENSE_AMT |

| DEATH |
|---|
| **PATID** |
| **DEATH_DATE** |
| DEATH_DATE_IMPUTE |
| **DEATH_SOURCE** |
| DEATH_MATCH_CONFIDENCE |

| DEATH_CAUSE |
|---|
| **PATID** |
| **DEATH_CAUSE** |
| **DEATH_CAUSE_CODE** |
| **DEATH_CAUSE_TYPE** |
| **DEATH_CAUSE_SOURCE** |
| DEATH_CAUSE_CONFIDENCE |

**Data captured from processes associated with healthcare delivery**

| VITAL |
|---|
| **VITALID** |
| **PATID** |
| ENCOUNTERID (optional) |
| **MEASURE_DATE** |
| MEASURE_TIME |
| **VITAL_SOURCE** |
| HT |
| WT |
| DIASTOLIC |
| SYSTOLIC |
| ORIGINAL_BMI |
| BP_POSITION |
| SMOKING |
| TOBACCO |
| TOBACCO_TYPE |

| CONDITION |
|---|
| **CONDITIONID** |
| **PATID** |
| ENCOUNTERID (optional) |
| REPORT_DATE |
| RESOLVE_DATE |
| ONSET_DATE |
| CONDITION_STATUS |
| **CONDITION** |
| **CONDITION_TYPE** |
| **CONDITION_SOURCE** |

| PRO_CM |
|---|
| **PRO_CM_ID** |
| **PATID** |
| ENCOUNTERID (optional) |
| **PRO_ITEM** |
| PRO_LOINC |
| **PRO_DATE** |
| PRO_TIME |
| **PRO_RESPONSE** |
| PRO_METHOD |
| PRO_MODE |
| PRO_CAT |

**Data captured within multiple contexts: healthcare delivery, registry activity, or directly from patients**

| ENCOUNTER |
|---|
| **ENCOUNTERID** |
| **PATID** |
| **ADMIT_DATE** |
| ADMIT_TIME |
| DISCHARGE_DATE |
| DISCHARGE_TIME |
| PROVIDERID |
| FACILITY_LOCATION |
| **ENC_TYPE** |
| FACILITYID |
| DISCHARGE_DISPOSITION |
| DISCHARGE_STATUS |
| DRG |
| DRG_TYPE |
| ADMITTING_SOURCE |

| DIAGNOSIS |
|---|
| **DIAGNOSISID** |
| **PATID** |
| **ENCOUNTERID** |
| *ENC_TYPE (replicated)* |
| *ADMIT_DATE (replicated)* |
| *PROVIDERID (replicated)* |
| **DX** |
| **DX_TYPE** |
| **DX_SOURCE** |
| PDX |

| PROCEDURES |
|---|
| **PROCEDURESID** |
| **PATID** |
| **ENCOUNTERID** |
| *ENC_TYPE (replicated)* |
| *ADMIT_DATE (replicated)* |
| *PROVIDERID (replicated)* |
| PX_DATE |
| **PX** |
| **PX_TYPE** |
| PX_SOURCE |

| LAB_RESULT_CM |
|---|
| **LAB_RESULT_CM_ID** |
| **PATID** |
| ENCOUNTERID (optional) |
| LAB_NAME |
| SPECIMEN_SOURCE |
| LAB_LOINC |
| PRIORITY |
| RESULT_LOC |
| LAB_PX |
| LAB_PX_TYPE |
| LAB_ORDER_DATE |
| SPECIMEN_DATE |
| SPECIMEN_TIME |
| **RESULT_DATE** |
| RESULT_TIME |
| RESULT_QUAL |
| RESULT_NUM |
| RESULT_MODIFIER |
| RESULT_UNIT |
| NORM_RANGE_LOW |
| NORM_MODIFIER_LOW |
| NORM_RANGE_HIGH |
| NORM_MODIFIER_HIGH |
| ABN_IND |

| PRESCRIBING |
|---|
| **PRESCRIBINGID** |
| **PATID** |
| ENCOUNTERID (optional) |
| RX_PROVIDERID |
| RX_ORDER_DATE |
| RX_ORDER_TIME |
| RX_START_DATE |
| RX_END_DATE |
| RX_QUANTITY |
| RX_REFILLS |
| RX_DAYS_SUPPLY |
| RX_FREQUENCY |
| RX_BASIS |
| RXNORM_CUI |

**Data captured from healthcare delivery, direct encounter basis**

| PCORNET_TRIAL |
|---|
| **PATID** |
| **TRIALID** |
| **PARTICIPANTID** |
| TRIAL_SITEID |
| TRIAL_ENROLL_DATE |
| TRIAL_END_DATE |
| TRIAL_WITHDRAW_DATE |
| TRIAL_INVITE_CODE |

**Associations with PCORnet clinical trials**

| HARVEST |
|---|
| **NETWORKID** |
| NETWORK_NAME |
| **DATAMARTID** |
| DATAMART_NAME |
| DATAMART_PLATFORM |
| CDM_VERSION |
| DATAMART_CLAIMS |
| DATAMART_EHR |
| BIRTH_DATE_MGMT |
| ENR_START_DATE_MGMT |
| ENR_END_DATE_MGMT |
| ADMIT_DATE_MGMT |
| DISCHARGE_DATE_MGMT |
| PX_DATE_MGMT |
| RX_ORDER_DATE_MGMT |
| RX_START_DATE_MGMT |
| RX_END_DATE_MGMT |
| DISPENSE_DATE_MGMT |
| LAB_ORDER_DATE_MGMT |
| SPECIMEN_DATE_MGMT |
| RESULT_DATE_MGMT |
| MEASURE_DATE_MGMT |
| ONSET_DATE_MGMT |
| REPORT_DATE_MGMT |
| RESOLVE_DATE_MGMT |
| PRO_DATE_MGMT |
| REFRESH_DEMOGRAPHIC_DATE |
| REFRESH_ENROLLMENT_DATE |
| REFRESH_ENCOUNTER_DATE |
| REFRESH_DIAGNOSIS_DATE |
| REFRESH_PROCEDURES_DATE |
| REFRESH_VITAL_DATE |
| REFRESH_DISPENSING_DATE |
| REFRESH_LAB_RESULT_CM_DATE |
| REFRESH_CONDITION_DATE |
| REFRESH_PRO_CM_DATE |
| REFRESH_PRESCRIBING_DATE |
| REFRESH_PCORNET_TRIAL_DATE |
| REFRESH_DEATH_DATE |
| REFRESH_DEATH_CAUSE_DATE |

**Process-related data**

http://www.pcornet.org/pcornet-common-data-model/

**Bold font** indicates fields that cannot be null due to primary key definitions or record-level constraints.

# LAB_RESULT_CM: contents

| LAB_RESULT_CM |
| --- |
| **LAB_RESULT_CM_ID** |
| **PATID** |
| ENCOUNTERID (optional) |
| LAB_NAME |
| SPECIMEN_SOURCE |
| LAB_LOINC |
| PRIORITY |
| RESULT_LOC |
| LAB_PX |
| LAB_PX_TYPE |
| LAB_ORDER_DATE |
| SPECIMEN_DATE |
| SPECIMEN_TIME |
| **RESULT_DATE** |
| RESULT_TIME |
| RESULT_QUAL |
| RESULT_NUM |
| RESULT_MODIFIER |
| RESULT_UNIT |
| NORM_RANGE_LOW |
| NORM_MODIFIER_LOW |
| NORM_RANGE_HIGH |
| NORM_MODIFIER_HIGH |
| ABN_IND |

## The “PCORnet 11”

Hemoglobin A1c
Creatine kinase total
Creatine kinase MB
Creatine kinase MB/creatine kinase total
Creatinine
Hemoglobin
Low-density lipoprotein
International normalized ratio
Troponin I cardiac
Troponin T cardiac (qualitative)
Troponin T cardiac (quantitative)

pcornet

AMIA

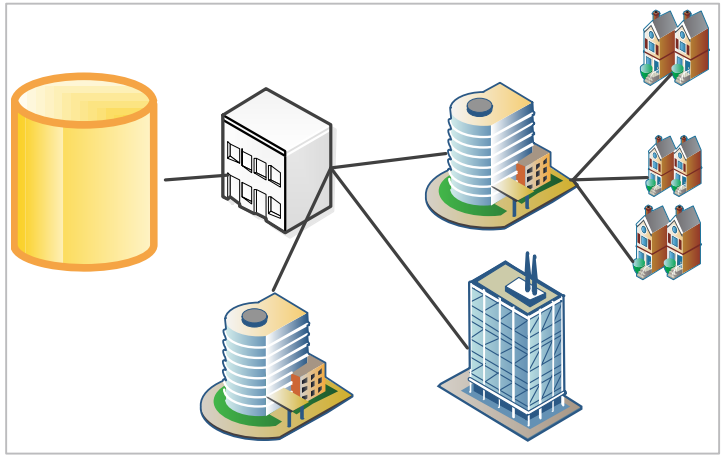
# Mapping labs in PCORnet is challenging: ***Diversity*** of data partners in PCORnet

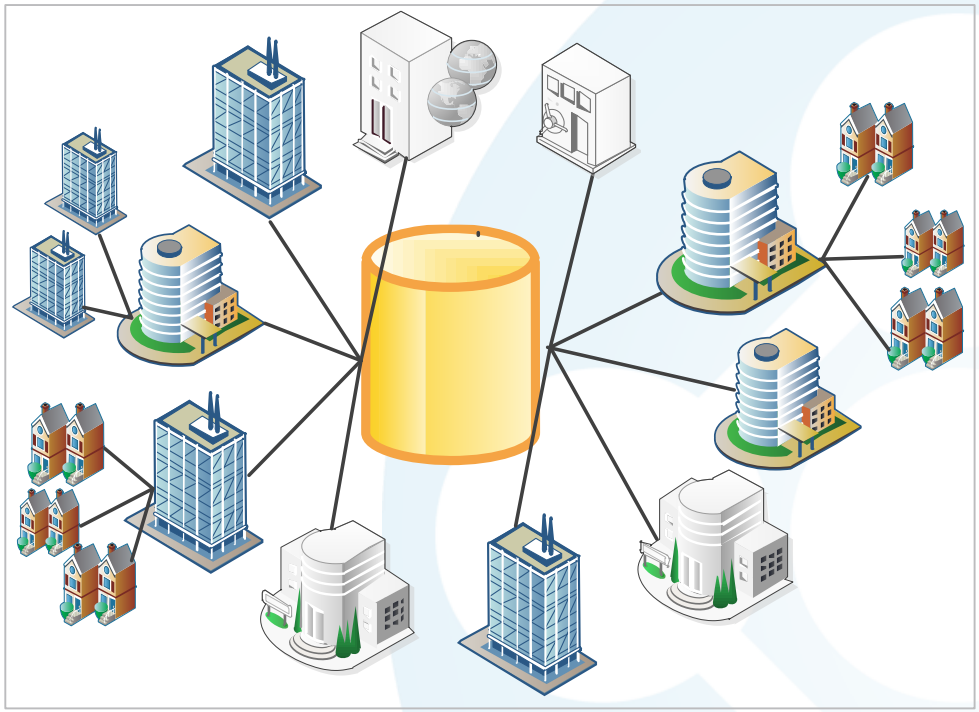
- 13 Clinical Data Research Networks + 20 Patient-Powered Research Networks
- ***Types*** of organizations in PCORnet
  - Children's hospital
  - Teaching hospital
  - Integrated health care delivery system
  - Public hospital
  - Academic medical center
  - Non-profit health care center/system
  - Not-for-profit health plan
  - Research hospital
  - Health information exchange (HIE)
  - Veterans Administration Medical Center

pcornet

AMIA

# Mapping labs in PCORnet is challenging: ***Composition*** of data partners in PCORnet

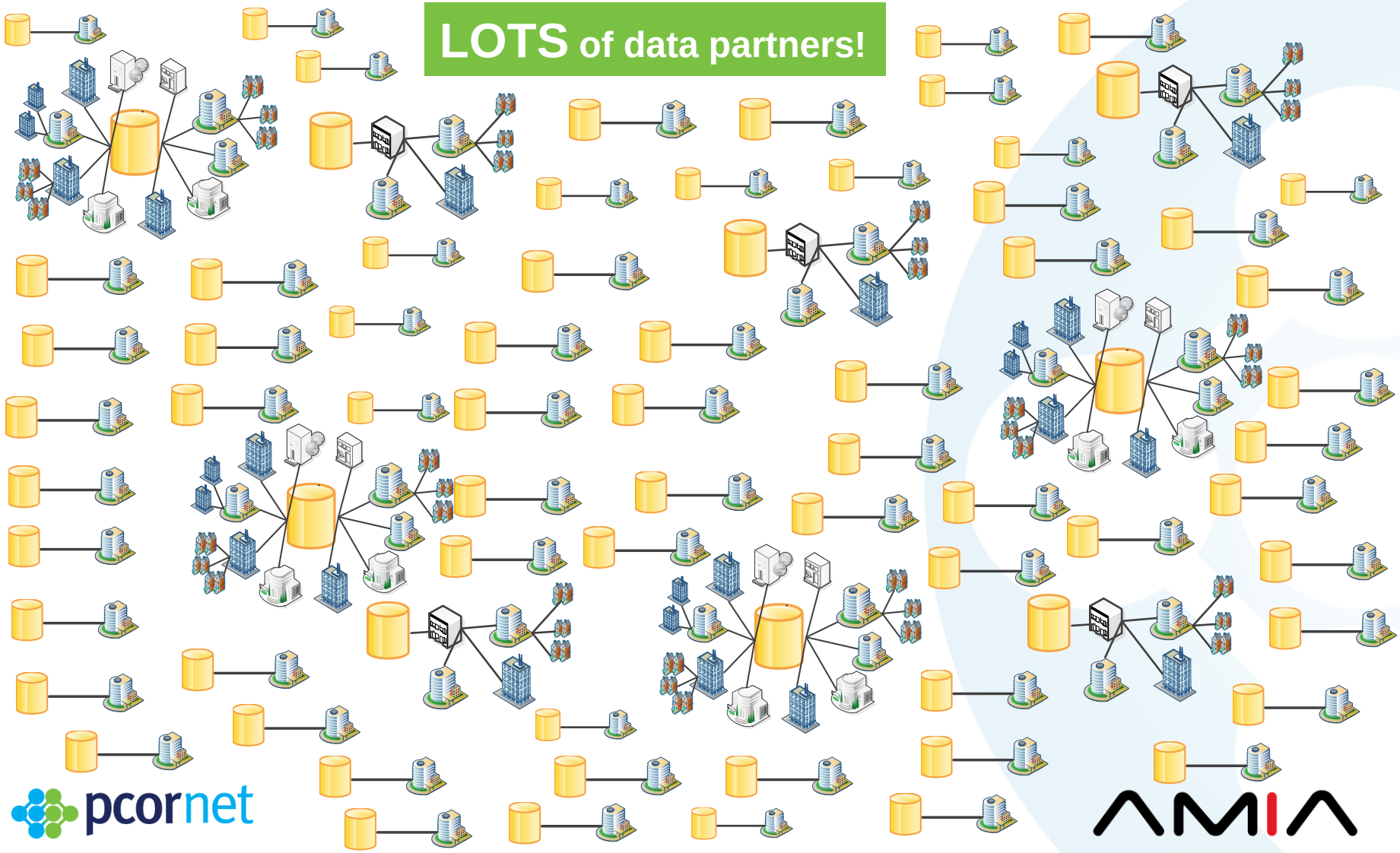




pcornet

AMIA

# Mapping labs in PCORnet is challenging: ***Density*** of data partners in PCORnet




pcornet

AMIA

# Mapping labs in PCORnet is challenging: ***State*** of source system lab data and mapping complexity

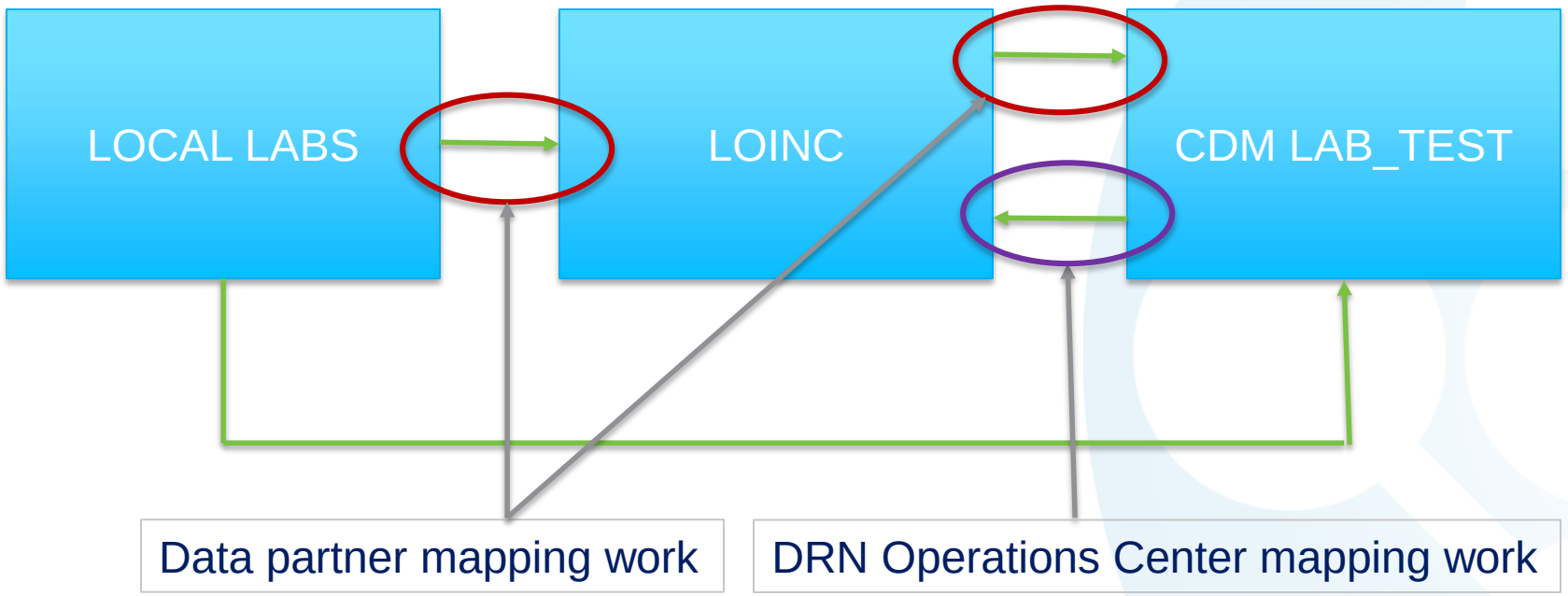
- Standardization varies
- Incomplete or idiosyncratic local names
- Multiple possible mappings
- Level of granularity choices

pcornet

AMIA

# Multiple Mapping Steps

Mapping from Local Labs to CDM “common measure” labs



# Engagement and Dissemination: CDM Forum

- Began in summer 2015 as an **outgrowth of the CDM Working Group**
- Purpose: Discuss CDM interpretation and implementation.
- Goal: **Connect people in different networks** who may have similar challenges.
- **Open sessions** ~ participants are generally **technical teams and implementation leadership**
- Good data partner engagement with an **average of 55 participants** per session

# CDM Forum Interest Groups

- Interest groups: **small engaged groups** formed around **specific topics** identified in the CDM Forum
- Agile complement to the larger CDM Forum
- Pilot interest group was **lab mappings**
  - **Purpose:** share challenges, information and experiences based on lab mapping activities
  - **Goal**: identify issues that can be tackled as a group…leverage the knowledge and expertise across PCORnet to learn from and help one another

# Lab Mapping Activity Survey

**PCORnet Lab Mapping Survey Plan**

**PURPOSE STATEMENT**

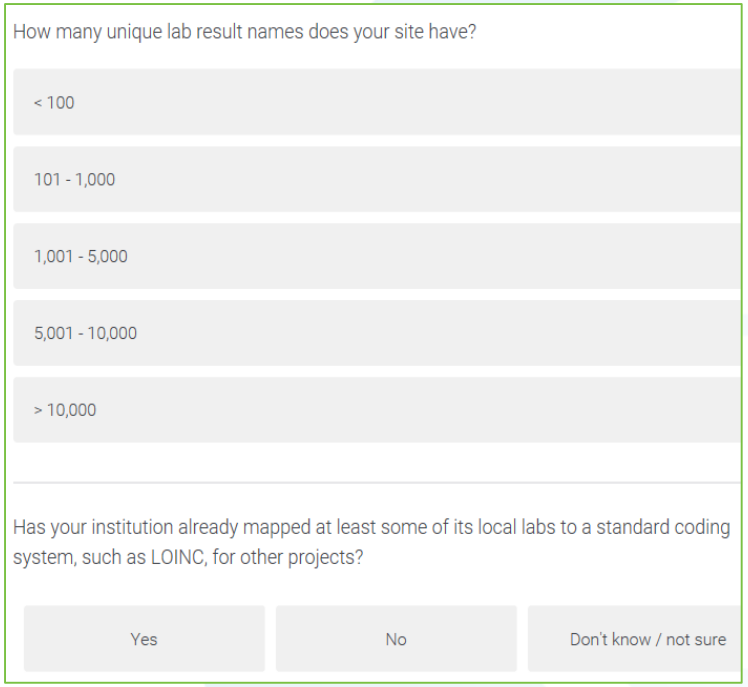
The goals of the PCORnet Lab Mapping Survey are to understand the landscape of lab mapping activity information from a greater number of sites/networks, and to utilize this information to recognize and support strategic approaches by sites/networks to map local system labs to the PCORnet CDM. This **voluntary** survey will build upon an AMIA abstract submitted earlier this fall describing the assessment of factors and approaches to lab mapping in PCORnet.

**DEVELOPMENT**

The PCORnet Lab Mapping Survey was developed collaboratively by members of the Lab Mapping Interest Group, under the leadership of Keith Marsolo. Survey design and activity coordination was led by Michelle Smerek, and Michael Park constructed the online survey in Qualtrics, based on specifications documented in MS Word.

**SCHEDULE**

- Nov. 11th – Survey opens, announced during CDM Implementation Forum. Survey invitation emails sent to:
  - Members of Lab Mapping Interest Group
  - CDM Implementation Forum mailing list
  - CDRN/DRNOC mailing list (Jenny Ibarra's list)
  - Posted to Central Desktop blog
  - Included in November 16TH Weekly Announcement




pcornet

AMIA

# Intriguing bits from survey results

34 total survey responses

12 different EHR systems reported

21 respondents have Epic

# Intriguing bits from survey results, cont.

87% of respondents:

Site has institution-specific lab codes

pcornet

AMIA

# More intriguing bits from survey results

How many unique lab result names does your site have?

24%: 1,001 - 5,000

32%: 5,001 - 10,000

36%: > 10,000 unique lab result names

# Intriguing bits from survey results, cont.

How many different synonyms does your site have for “hemoglobin?”

64%: > 7 synonyms for “hemoglobin”

# Intriguing bits from survey results, cont.

85% of respondents

…already mapped at least some local labs to a standard coding system, such as LOINC

# Intriguing bits from survey results, cont.

What % of total lab results (by volume) did the mapping cover?

9%: 100% covered

26%: 75-99% covered

13%: 50-75% covered

17%: 25-50% covered

9%: < 25% covered

pcornet

AMIA

# Successful approaches for addressing challenges

- **Pareto principle**
  - A small number of local tests account for the majority of total test volume
  - Employ an 80/20 rule to target high frequency labs for mapping
- **Align quality criteria with use case**
  - Don’t let perfection be the enemy of the good
  - Define “appropriateness for use” criteria
  - Set a maximum level for acceptable missing values

# Successful approaches cont.

- **Cross-functional mapping team**
  - Technical analysts working with a clinical lab subject matter expert to determine best mappings.
  - Engaging clinical teams allows analysts to provide feedback that can be used to increase standardization and data quality upstream.
- **Map ⇨ review ⇨ repeat**
  - Work in an iterative fashion: group subsets of lab results in the source data, map, review with lab domain experts, then refine.

# Summary

## Approaching lab mapping in PCORnet from multiple angles

- CDM Implementation Forum and Lab Mapping Interest Group
  - Sharing challenges, strategies, processes, and lessons learned
- Leveraging lab data standards
  - Best Practices Sharing session by LOINC expert Daniel Vreeman
  - Expanded CDM reference table that includes LOINC codes that correspond to LAB_NAME common measures
- Suggested vs. prescriptive guidance
  - With such a large, diverse group of data partners, there is no one-size-fits-all solution
  - **Gathering** and **sharing** information that will help the people doing the mapping work

# Further Reading


Raebel MA, Haynes K, Woodworth TS, et al. Electronic clinical laboratory test results data tables: lessons from Mini-Sentinel. Pharmacoepidemiol Drug Saf. 2014 Jun; 23(6):609-18.

Priest EL, Klekar C, Cantu G, Berryman C, Garinger G, Hall L, Kouznetsova M, Kudyakov R, Masica A. Developingelectronic data methods infrastructure to participate in collaborative research networks. EGEMS (Wash DC). 2014 Dec 2;2(1):1126.8.

Bayley KB, Belnap T, Savitz L, Masica AL, Shah N, Fleming NS. Challenges in using electronic health record data for CER: experience of 4 learning organizations and solutions applied. Med Care. 2013 Aug;51(8 Suppl 3):S80-6.

Vreeman DJ, Hook J, Dixon BE. Learning from the crowd while mapping to LOINC. J Am Med Inform Assoc. 2015 Nov;22(6):1205-11

Khan AN, Griffith SP, Moore C, et al. Standardizing laboratory data by mapping to LOINC. J Am Med Inform Assoc. 2006 May-Jun;13(3):353-5.

McDonald CJ, Huff SM, Suico JG, et al. LOINC, a universal standard for identifying laboratory observations: a 5- year update. Clin Chem 2003;49;624-33.

Dixon BE, Hook J, Vreeman DJ. Learning from the crowd in terminology mapping: the LOINC experience. Lab Med 2015 Spring;46(2):168-174.

Logical Object Identifiers Names and Codes (LOINC)® Available at: https://loinc.org/. Accessed Sept 21, 2015.

Lin MC, Vreeman DJ, McDonald CJ, Huff SM. A Characterization of Local LOINC Mapping for Laboratory Tests in Three Large Institutions. Methods Inf Med. 2011; 50(2):105-14.

# Acknowledgements

"No one can whistle a symphony; it takes a whole orchestra to play it."

*~H.E. Luccock*

We gratefully acknowledge the participation of the following Lab Mapping Enthusiasts

**Lab Mapping Interest Group**

Rimma Belenkaya, Angela Bos, Rob Follett, W. Ed Hammond, Rahul Jain, Pedro Rivera, Thom Mathes, Allison O'Neill, Eliel Olivera, James Topping, Michael Park, Elisa Priest, Claudia Pulgarin, Jon Puro, Deepa Ranka, Iben Ricket, Thuy Le, Trent Rosenbloom, Toan Ong, Shelley Rusincovitch, Anthony Solomonides

**PCORnet Key Thought Leadership**

Jeff Brown, Lesley Curtis, Daniella Meeker, Abel Kho, Keith Marsolo

**We thank members of Regenstrief for their generous support and participation**

Swapna Abhyankar, Sara Armson, Clem McDonald, Theresa Pritchard, Daniel Vreeman

# Funding

This work was funded by the Patient-Centered Outcomes Research Institute (PCORI) (PCO-COORDCTR2013) for development of the National Patient-Centered Clinical Research Network, known as PCORnet.

The views presented in this work are solely the responsibility of the author(s) and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee or other participants in PCORnet.

# Questions?

# Thank You!

**Michelle M. Smerek**
Research Informaticist
Clinical Research Informatics
Duke Clinical Research Institute

Michelle.smerek@duke.edu
Phone: 919-668-8910