PCORnet CDM Forum

Wednesday, June 8, 2016 from 2–3 PM Eastern
Hosted by Keith Marsolo, PhD
Facilitated by Shelley Rusincovitch and Michelle Smerek

pcornet℠ | The National Patient-Centered Clinical Research Network

Agenda

- Welcome, announcements, and brief review of issue tracker
- Implementation issues identified through data characterization
- Exploratory analysis of variation in procedure data
- Read-only database accounts
- Wrap up

pcornet

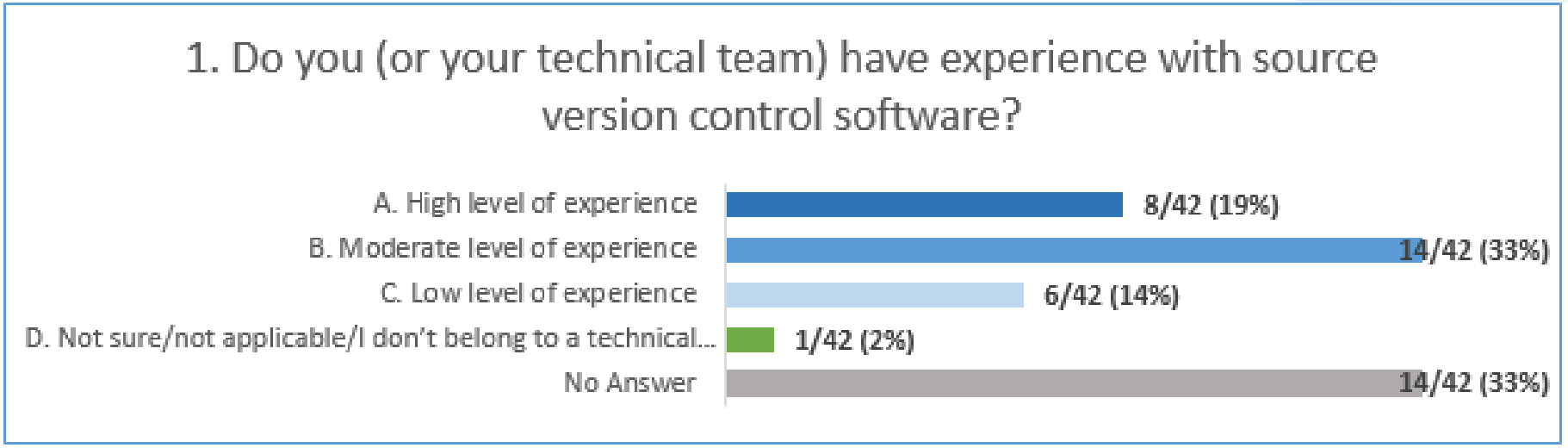
Announcements

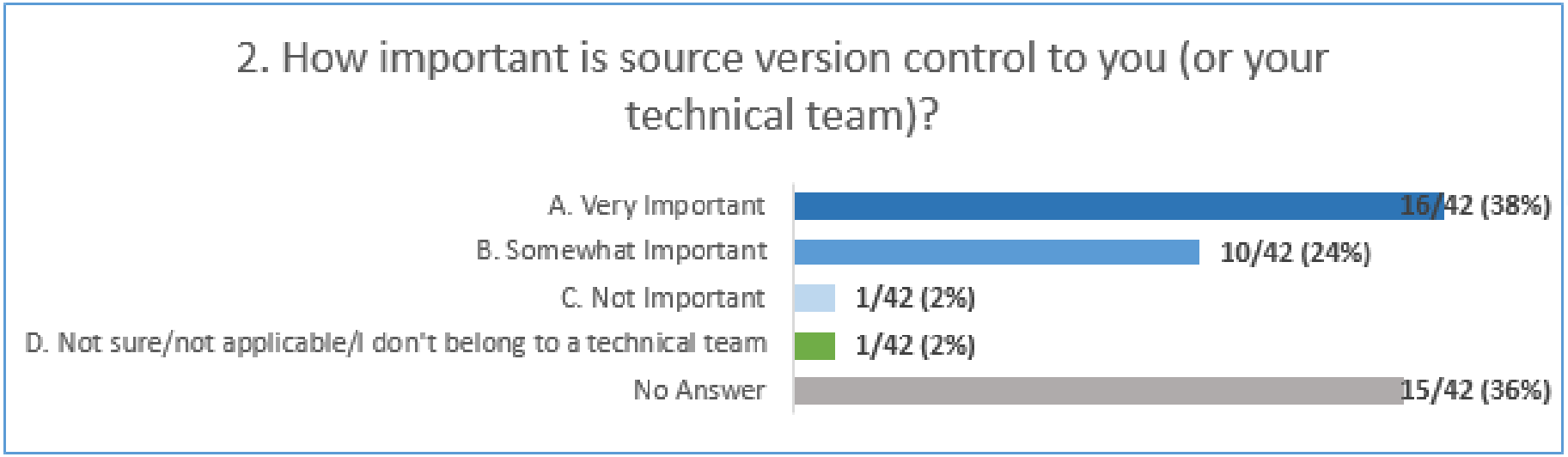
Recap: CDM Forum from May 11

Presentation by Michael Matheny on version control software and repository resources

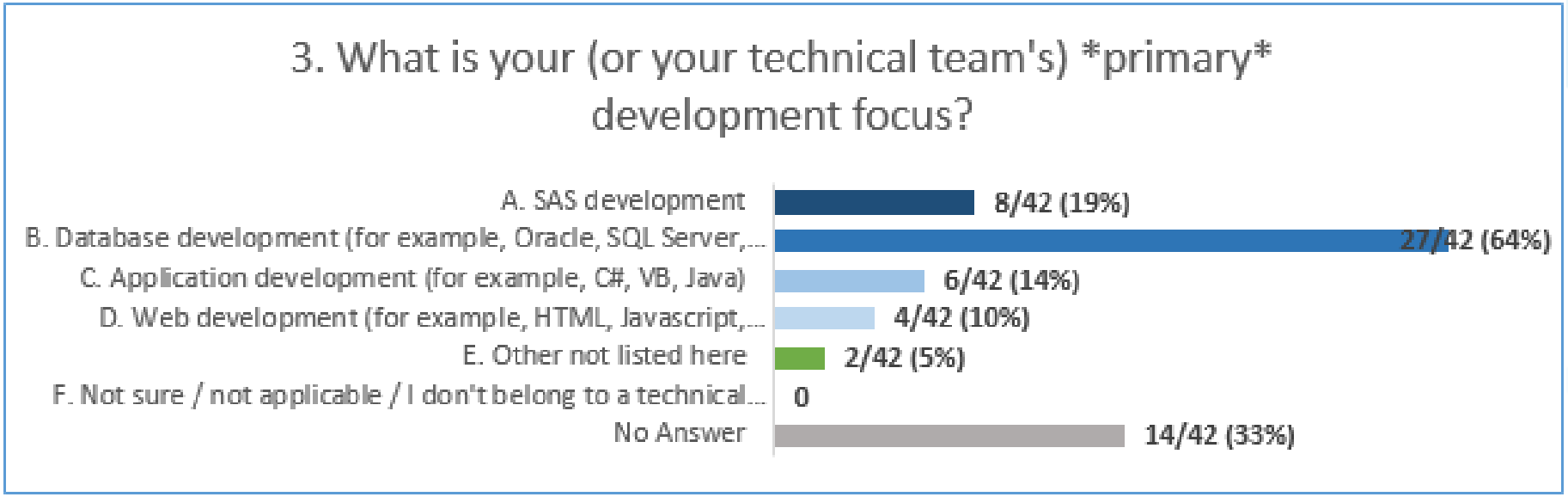
- Slides: https://github.com/CDMFORUM/CDM-GUIDANCE/wiki/CDM-Forum-Materials
- Recording and meeting summary: https://pcornet.imeetcentral.com/p/ZgAAAAAAdJMO

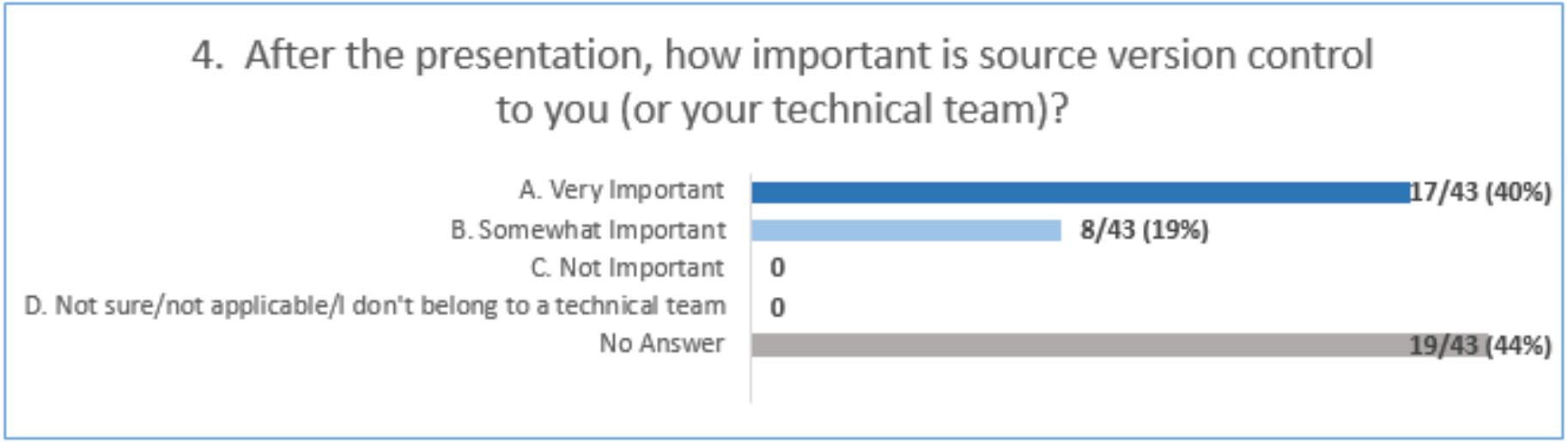
Poll Results from CDM Implementation Forum on May 11, 2016





Poll Results from CDM Implementation Forum on May 11, 2016 (continued)





Recap: Research Using a DRN Forum from May 31

Presentation by Jeff Brown on menu-driven querying in PCORnet

- Slides and meeting recording: https://pcornet.imeetcentral.com/p/ZgAAAAAAdYBr

Recap: DRN OC Forum from June 6

“Approved for Research and Approved for PTR: What Next?”

- Slides and meeting recording: https://pcornet.imeetcentral.com/p/ZgAAAAAAdcWu

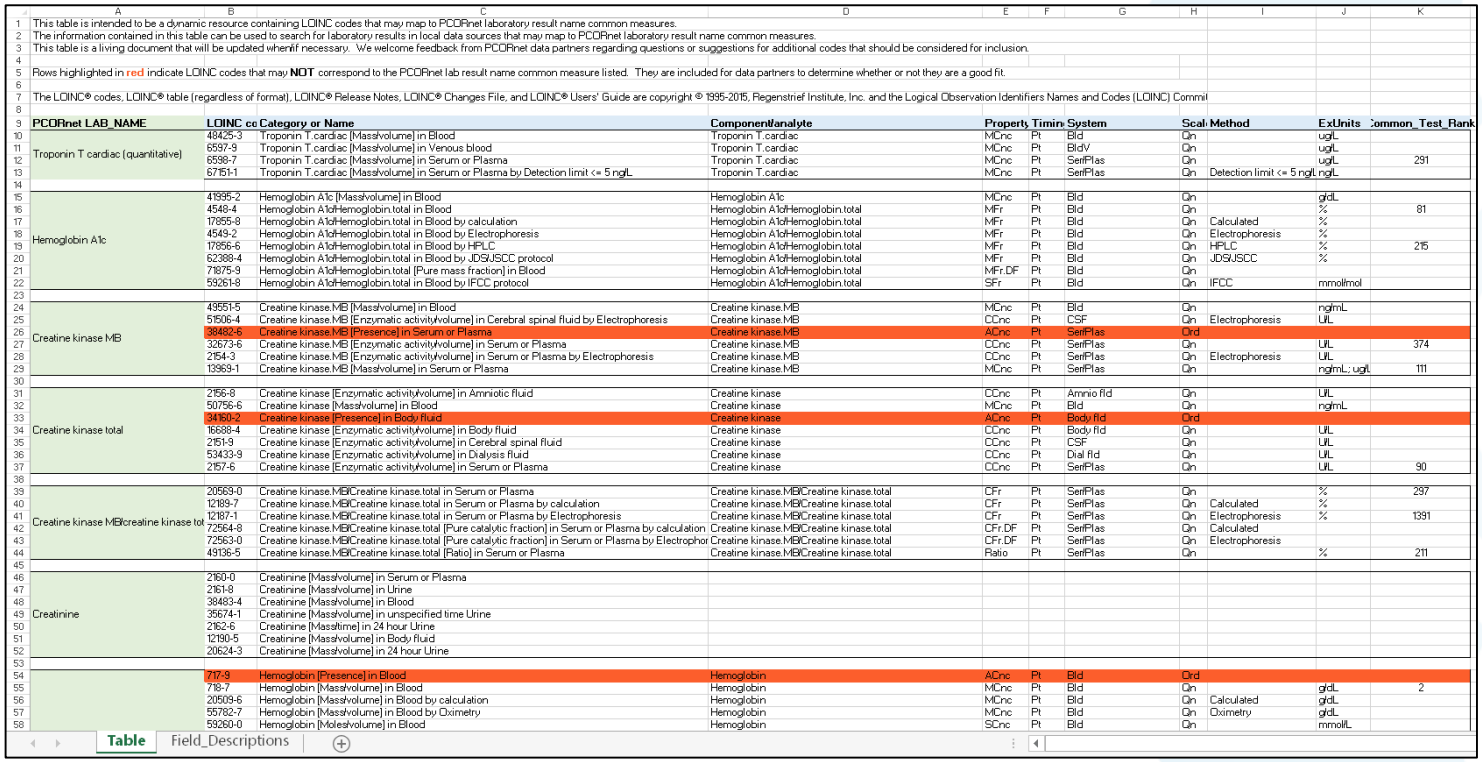
(PTR = “Preparatory to Research”)

Interest Group Updates

- Medication Mapping Interest Group had a call May 16th.
  - Recording posted to iMeetCentral: https://pcornet.imeetcentral.com/p/aQAAAAACuUu1
- Lab Mapping Interest Group had a call May 6th.
  - Recording posted to iMeetCentral: https://pcornet.imeetcentral.com/p/aQAAAAACuFPd
  - Updated lab result reference table posted to GitHub: https://github.com/CDMFORUM/CDM-GUIDANCE/wiki/Lab-Mapping-Resources

Lab Mappings

- Updated LOINC Common Measures table uploaded in GitHub, with thanks to the Lab Mappings interest group for their feedback!



This table is intended to be a dynamic resource containing LOINC codes that may map to PCORnet laboratory result name common measures.
The information contained in this table can be used to search for laboratory results in local data sources that may map to PCORnet laboratory result name common measures.
This table is a living document that will be updated when/if necessary. We welcome feedback from PCORnet data partners regarding questions or suggestions for additional codes that should be considered for inclusion.

Rows highlighted in red indicate LOINC codes that may **NOT** correspond to the PCORnet lab result name common measure listed. They are included for data partners to determine whether or not they are a good fit.

The LOINC® codes, LOINC® table (regardless of format), LOINC® Release Notes, LOINC® Changes File, and LOINC® Users' Guide are copyright © 1995-2015, Regenstrief Institute, Inc. and the Logical Observation Identifiers Names and Codes (LOINC) Commit

PCORnet LAB_NAME	LOINC co	Category or Name	Component/analyte	Property	Timin	System	Scal	Method	ExUnits	Common_Test_Rank
Troponin T cardiac (quantitative)	48425-3	Troponin T.cardiac [Mass/volume] in Blood	Troponin T.cardiac	MCnc	Pt	Bld	Qn		ug/L	
	6597-9	Troponin T.cardiac [Mass/volume] in Venous blood	Troponin T.cardiac	MCnc	Pt	BldV	Qn		ug/L	
	6598-7	Troponin T.cardiac [Mass/volume] in Serum or Plasma	Troponin T.cardiac	MCnc	Pt	SerPlas	Qn		ug/L	291
	67151-1	Troponin T.cardiac [Mass/volume] in Serum or Plasma by Detection limit <= 5 ng/L	Troponin T.cardiac	MCnc	Pt	SerPlas	Qn	Detection limit <= 5 ng/L	ng/L	
Hemoglobin A1c	41995-2	Hemoglobin A1c [Mass/volume] in Blood	Hemoglobin A1c	MCnc	Pt	Bld	Qn		g/dL	
	4548-4	Hemoglobin A1c/Hemoglobin.total in Blood	Hemoglobin A1c/Hemoglobin.total	MFr	Pt	Bld	Qn		%	81
	17855-8	Hemoglobin A1c/Hemoglobin.total in Blood by calculation	Hemoglobin A1c/Hemoglobin.total	MFr	Pt	Bld	Qn	Calculated	%	
	4549-2	Hemoglobin A1c/Hemoglobin.total in Blood by Electrophoresis	Hemoglobin A1c/Hemoglobin.total	MFr	Pt	Bld	Qn	Electrophoresis	%	
	17856-6	Hemoglobin A1c/Hemoglobin.total in Blood by HPLC	Hemoglobin A1c/Hemoglobin.total	MFr	Pt	Bld	Qn	HPLC	%	215
	62388-4	Hemoglobin A1c/Hemoglobin.total in Blood by JDS/JSCC protocol	Hemoglobin A1c/Hemoglobin.total	MFr	Pt	Bld	Qn	JDS/JSCC	%	
	71875-9	Hemoglobin A1c/Hemoglobin.total [Pure mass fraction] in Blood	Hemoglobin A1c/Hemoglobin.total	MFr.DF	Pt	Bld	Qn			
	59261-8	Hemoglobin A1c/Hemoglobin.total in Blood by IFCC protocol	Hemoglobin A1c/Hemoglobin.total	SFr	Pt	Bld	Qn	IFCC	mmol/mol	
Creatine kinase MB	49551-5	Creatine kinase.MB [Mass/volume] in Blood	Creatine kinase.MB	MCnc	Pt	Bld	Qn		ng/mL	
	51506-4	Creatine kinase.MB [Enzymatic activity/volume] in Cerebral spinal fluid by Electrophoresis	Creatine kinase.MB	CCnc	Pt	CSF	Qn	Electrophoresis	U/L	
	38482-6	Creatine kinase.MB [Presence] in Serum or Plasma	Creatine kinase.MB	ACnc	Pt	SerPlas	Ord			
	32673-6	Creatine kinase.MB [Enzymatic activity/volume] in Serum or Plasma	Creatine kinase.MB	CCnc	Pt	SerPlas	Qn		U/L	374
	2154-3	Creatine kinase.MB [Enzymatic activity/volume] in Serum or Plasma by Electrophoresis	Creatine kinase.MB	CCnc	Pt	SerPlas	Qn	Electrophoresis	U/L	
	13969-1	Creatine kinase.MB [Mass/volume] in Serum or Plasma	Creatine kinase.MB	MCnc	Pt	SerPlas	Qn		ng/mL; ug/L	111
Creatine kinase total	2156-8	Creatine kinase [Enzymatic activity/volume] in Amniotic fluid	Creatine kinase	CCnc	Pt	Amnio fld	Qn		U/L	
	50756-6	Creatine kinase [Mass/volume] in Blood	Creatine kinase	MCnc	Pt	Bld	Qn		ng/mL	
	34150-2	Creatine kinase [Presence] in Body fluid	Creatine kinase	ACnc	Pt	Body fld	Ord			
	16688-4	Creatine kinase [Enzymatic activity/volume] in Body fluid	Creatine kinase	CCnc	Pt	Body fld	Qn		U/L	
	2151-9	Creatine kinase [Enzymatic activity/volume] in Cerebral spinal fluid	Creatine kinase	CCnc	Pt	CSF	Qn		U/L	
	53433-9	Creatine kinase [Enzymatic activity/volume] in Dialysis fluid	Creatine kinase	CCnc	Pt	Dial fld	Qn		U/L	
	2157-6	Creatine kinase [Enzymatic activity/volume] in Serum or Plasma	Creatine kinase	CCnc	Pt	SerPlas	Qn		U/L	90
Creatine kinase MB/creatine kinase tot	20569-0	Creatine kinase.MB/Creatine kinase.total in Serum or Plasma	Creatine kinase.MB/Creatine kinase.total	CFr	Pt	SerPlas	Qn		%	297
	12189-7	Creatine kinase.MB/Creatine kinase.total in Serum or Plasma by calculation	Creatine kinase.MB/Creatine kinase.total	CFr	Pt	SerPlas	Qn	Calculated	%	
	12187-1	Creatine kinase.MB/Creatine kinase.total in Serum or Plasma by Electrophoresis	Creatine kinase.MB/Creatine kinase.total	CFr	Pt	SerPlas	Qn	Electrophoresis	%	1391
	72564-8	Creatine kinase.MB/Creatine kinase.total [Pure catalytic fraction] in Serum or Plasma by calculation	Creatine kinase.MB/Creatine kinase.total	CFr.DF	Pt	SerPlas	Qn	Calculated		
	72563-0	Creatine kinase.MB/Creatine kinase.total [Pure catalytic fraction] in Serum or Plasma by Electrophor	Creatine kinase.MB/Creatine kinase.total	CFr.DF	Pt	SerPlas	Qn	Electrophoresis		
	49136-5	Creatine kinase.MB/Creatine kinase.total [Ratio] in Serum or Plasma	Creatine kinase.MB/Creatine kinase.total	Ratio	Pt	SerPlas	Qn		%	211
Creatinine	2160-0	Creatinine [Mass/volume] in Serum or Plasma								
	2161-8	Creatinine [Mass/volume] in Urine								
	38483-4	Creatinine [Mass/volume] in Blood								
	35674-1	Creatinine [Mass/volume] in unspecified time Urine								
	2162-6	Creatinine [Mass/time] in 24 hour Urine								
	12190-5	Creatinine [Mass/volume] in Body fluid								
	20624-3	Creatinine [Mass/volume] in 24 hour Urine								
	717-9	Hemoglobin [Presence] in Blood	Hemoglobin	ACnc	Pt	Bld	Ord			
	718-7	Hemoglobin [Mass/volume] in Blood	Hemoglobin	MCnc	Pt	Bld	Qn		g/dL	2
	20509-6	Hemoglobin [Mass/volume] in Blood by calculation	Hemoglobin	MCnc	Pt	Bld	Qn	Calculated	g/dL	
	55782-7	Hemoglobin [Mass/volume] in Blood by Oximetry	Hemoglobin	MCnc	Pt	Bld	Qn	Oximetry	g/dL	
	59260-0	Hemoglobin [Moles/volume] in Blood	Hemoglobin	SCnc	Pt	Bld	Qn		mmol/L	

Table | Field_Descriptions

https://github.com/CDMFORUM/CDM-GUIDANCE/wiki/Lab-Mapping-Resources

Review of Issue Tracker (live)

CDM errata issue tracker:
https://github.com/CDMFORUM/CDM-ERRATA/issues

CDM guidance issue tracker:
https://github.com/CDMFORUM/CDM-GUIDANCE/issues

Implementation issues identified through data characterization

Keith Marsolo, PhD

pcornet

Research in a distributed research network

- How do you ask a research question at hundreds of institutions and get back results you can trust?
  - Option 1 — Write a description and have everyone create a local implementation to run on their data
  - **Option 2 — Create an algorithm that can run against a single, common data model**

pcornet

Benefit of a Common Data Model

Same data are represented differently at different institutions
(e.g., Type of Encounter)

SITE 1
Social Work Visit
Allied Health
Office Visit
Nurse Visit
Procedure Visit
Employee Health
Vascular Lab
Sleep Study Visit
Social Work Visit

SITE 2
Office Visit
Specimen
Postpartum Visit
Clinical Support
Initial Prenatal

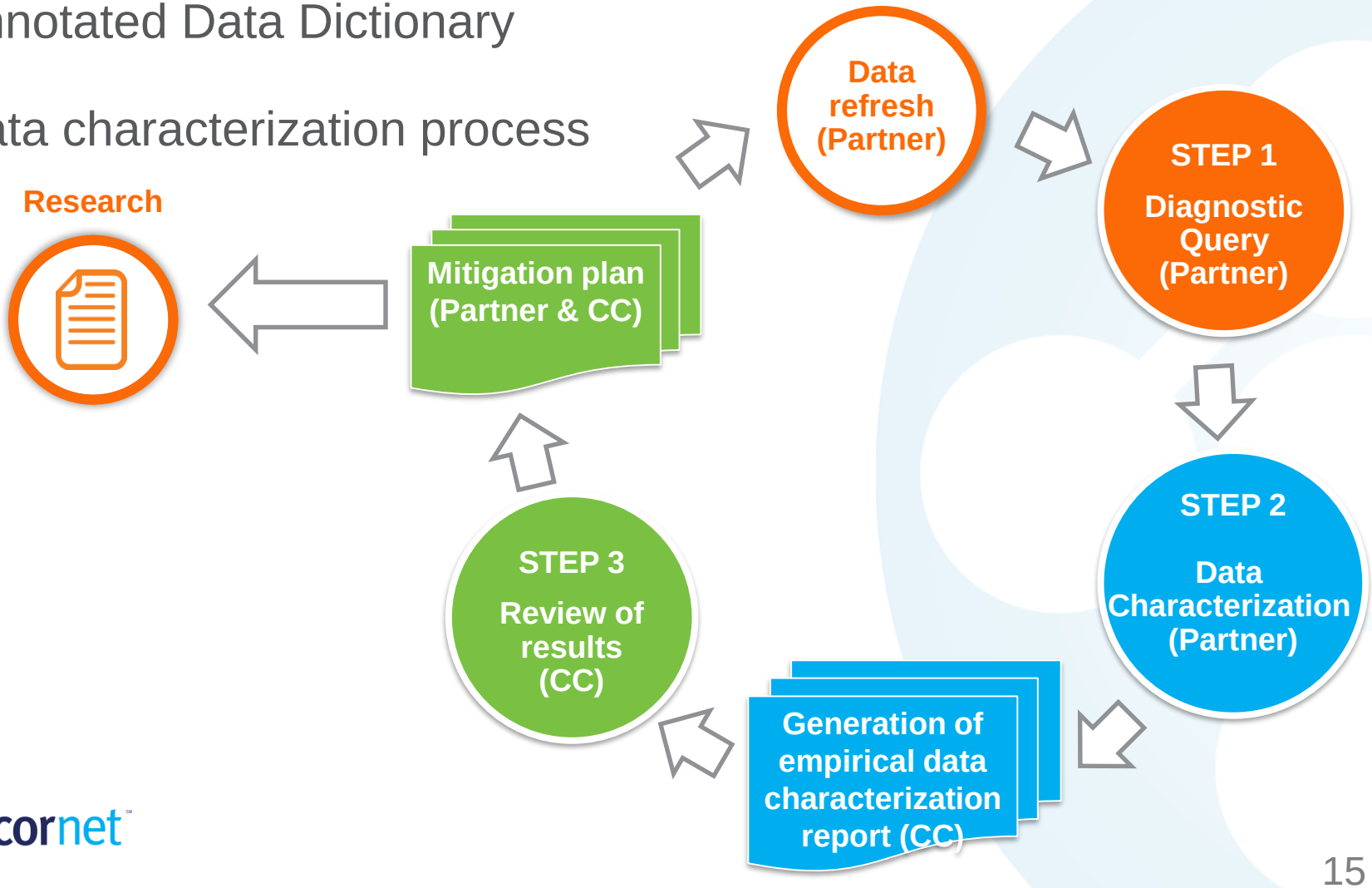
SITE 3
Home Care Visit
Office Visit
Therapy Visit
Orders Only
Cardiology Testing
Hospital Encounter

Common Data Model
Ambulatory Visit (AV)
Emergency Department (ED)
ED Admit to Inpatient (EI)
Inpatient Hospital (IP)
Non-Acute Inst. Stay (IS)
Other Ambulatory (OA)
Other (OT)
Unknown (UN)
No Information (NI)
(null)
Ambulatory Visit (AV)

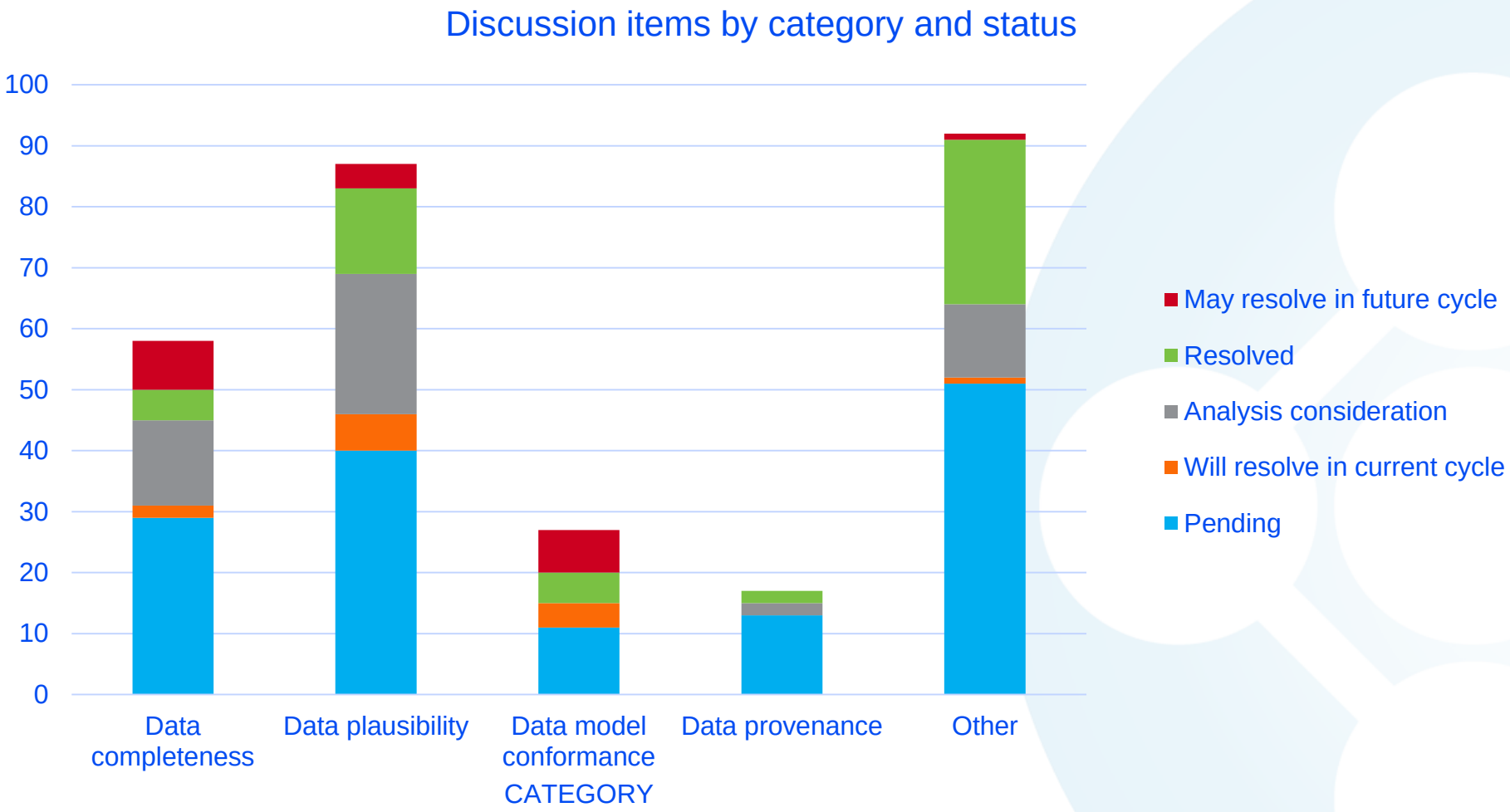
In order to be able to trust results of an analysis, need to have *consistent representations*

How do we assess whether we have consistent representations?

- Annotated Data Dictionary
- Data characterization process



Data Characterization Reviews
Classification of discussion items



As of June 6

Some emerging themes…

- ENCOUNTER
  - Heterogeneity in defining an encounter, but most are creating 1 encounter per patient per encounter type per provider per day
- DIAGNOSIS and PROCEDURES
  - Heterogeneity in data provenance (professional billing, facility/technical billing, claims, physician EHR, and/or facility EHR)
- PROCEDURES table
  - PX_TYPE: Predominantly ICD and CPT
  - Most DataMarts are capturing laboratory orders (CPT codes)
  - **CPT/HCPCS codes may not be assigned to the correct CPT/HCPCS subtype**
  - **Heterogeneity in PX codes within a given PX_TYPE**
- Impact of changes in source data or systems
  - Example: all patients with crossover encounters discharged from one EHR and admitted to the new EHR

Discussion of implementation issues in CDM Forums

- To date:
  - Focus on questions/topics of interest identified by data partners
  - Form interest groups to further discuss and/or identify best practices
- Going forward:
  - Focus on issues identified through data characterization
  - Gather additional information from partners; identify & document potential resolution(s); communicate findings (process TBD)

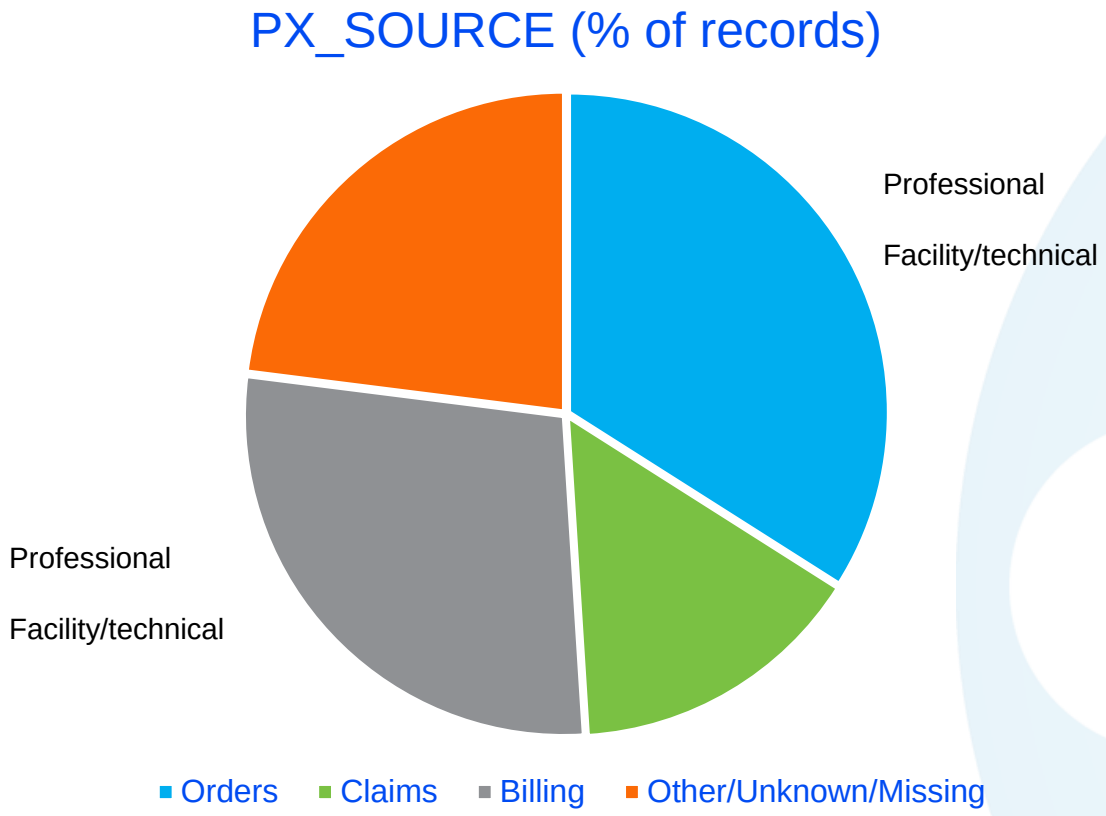
pcornet

Exploratory analysis of variation in procedure data

Laura Qualls, MS

pcornet

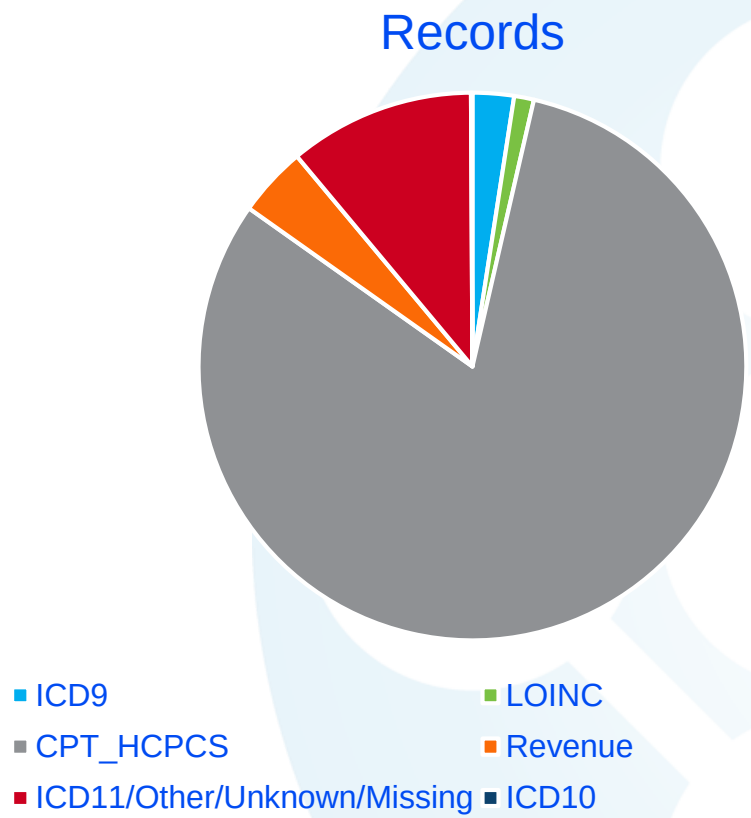
PROCEDURES Overview
Data Provenance



PROCEDURES Overview
PX_TYPES

- ICD
  - 09=ICD-9-CM
  - 10=ICD-10-PCS*
  - 11=ICD-11-PCS
- CPT/HCPCS
  - C2=CPT Category II
  - C3=CPT Category III
  - C4=CPT Category I**

  } HCPCS Level I
  - HC=HCPCS (i.e., HCPCS Level II)
  - H3=HCPCS Level III
- Other terminologies
  - Revenue (RE)
  - NDC (ND)
  - LC (LOINC)
- Flavors of null (NI/UN/OT/Null or Missing)



* Required as of Oct 2015. Most DataMarts's currently include data no later than early 2016.
**https://github.com/CDMFORUM/CDM-ERRATA/issues/14

Exploratory analyses

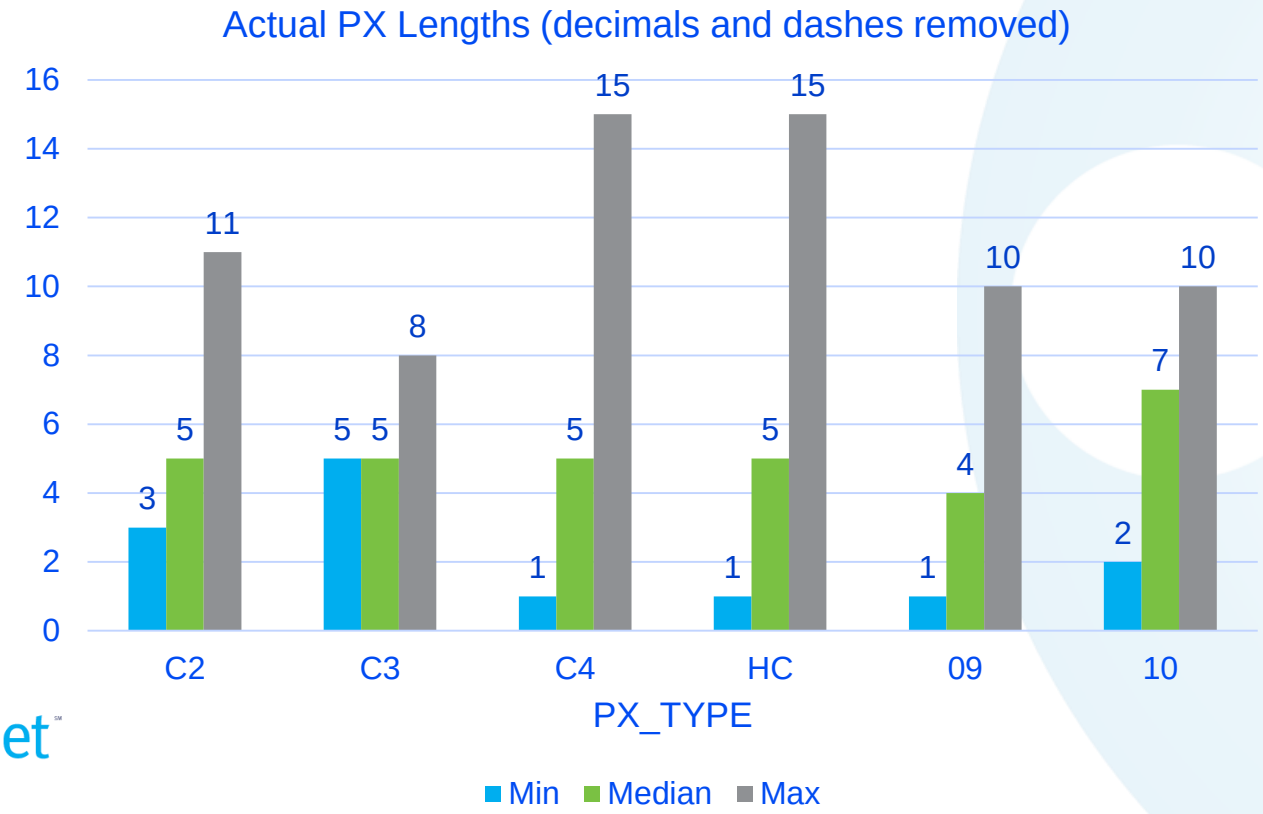
- Q1: Do codes meet the basic format for the given code type?*
  - Assumption: Computable phenotypes will incorporate PX_TYPE
- Q2: Are selected CPT and HCPCS codes accurately mapped to the correct PX_TYPE?
- Q3: Are there signals of potentially incomplete data capture?

*Data characterization does not currently include cross-referencing codes against reference tables

pcornet

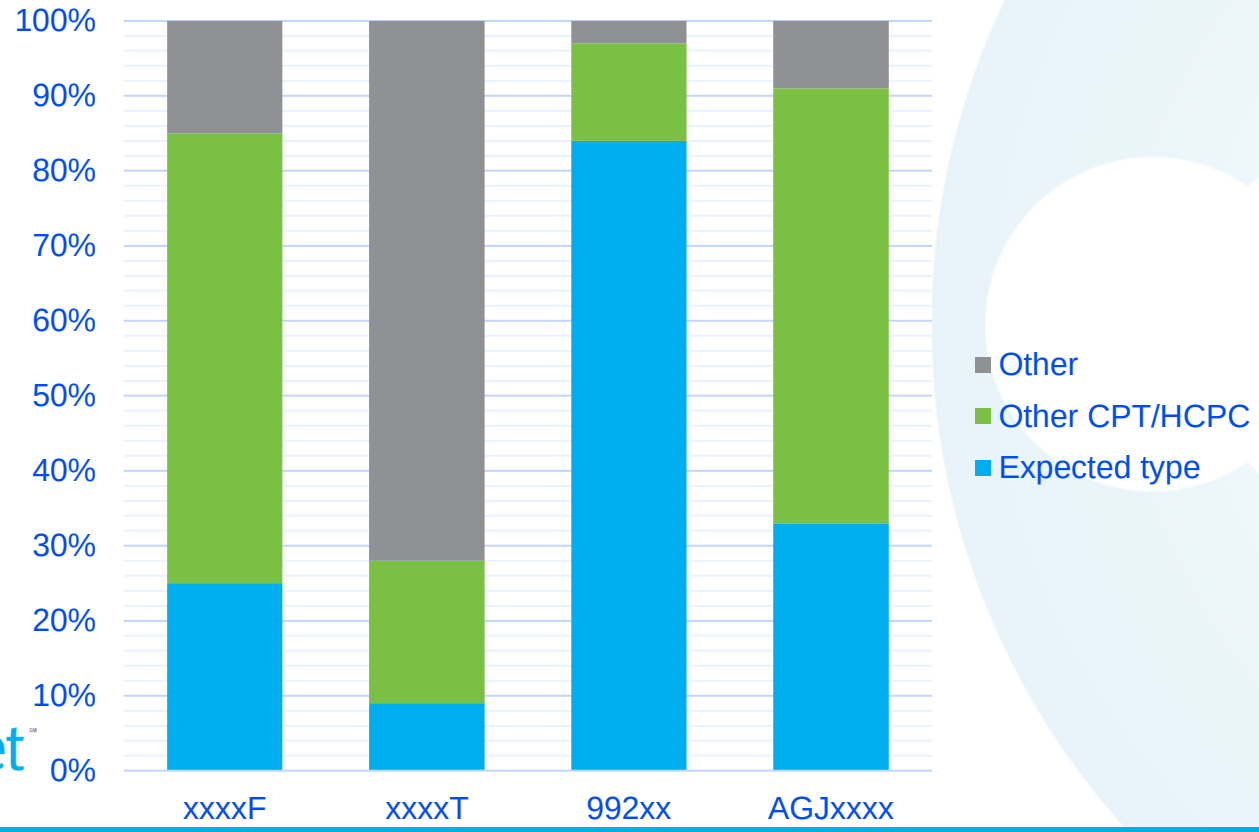
Q1. Do codes meet the basic format for the given procedure type?

- Basic format
  - ICD-9-CM (09) : 3-4 numbers (e.g. 89.39)
  - ICD-10-PCS (10): 7 alphanumeric characters (e.g. 09B00ZX).
  - CPT/HCPCS (C2/C3/C4/HC/H3): 5 alphanumeric characters (e.g. H2010, 99213); may be longer if modifiers are included



Q2. Are selected CPT and HCPCS codes mapped to the correct procedure type?

- Sample codes
  - xxxxF codes (codes used for tracking performance measurement) should be C2
  - xxxxT codes (temporary codes for emerging technologies, services, and procedures) should be C3
  - 992xx codes (basic E/M codes) should be C4.
  - A/G/Jxxxx (HCPCS Level II codes for various services not covered by CPT) should be HC.



Q3. Are there signals of potentially incomplete data capture?

A. We identified the top 5 most common procedures codes (ICD9, CPT lab orders, other CPT codes, and HCPCS) across PCORnet and looked to see if they were present in each DataMart
   - For most DataMarts, if the code type was used, most or all of the top 5 codes were present
   - This was less likely to be the case for lab test orders (8xxxx), which were more likely to either be missing entirely, and/or present but not to include the most common codes (e.g. 85025 complete blood count)

B. We observed that in approximately 20% of DataMarts, more than half of the patients with encounters did not have any procedure data*
   - Potential omissions: data sources (professional vs. facility, orders vs. billing); procedures documented at the detail/line level; procedures without hard charges; lab test orders; others?

*Procedure data by encounter type and over time is shown in the Empirical Data Characterization Report, Table and Chart IVB

pcornet

Topics for Discussion

- PX_TYPES
  - How do you ensure that codes are mapped to the correct PX_TYPE?
  - How do you map local codes (order data) to standard terminologies?
- How do you detect potentially incomplete capture of procedure data?
- Other topics?

Read-only Database Accounts

All

pcornet

The question

- “Can queries be run using a read-only database account?”
  - Good practice in security measures
- Data partner experience: read-only account presented no difficulty

pcornet

Beginning with a misunderstanding…

```
6169
6170  %macro xminmax(idsn);
6171
6172  * append each variable into base dataset *;
6173  proc append base=query data=&idsn;
6174  run;
6175
6176  %mend xminmax;
6177  %xminmax(idsn=demographic_birth_date)
6178  %xminmax(idsn=encounter_admit_date)
6179  %xminmax(idsn=encounter_discharge_date)
6180  %xminmax(idsn=diagnosis_admit_date)
6181  %xminmax(idsn=procedures_admit_date)
6182  %xminmax(idsn=procedures_px_date)
6183  %xminmax(idsn=vital_measure_date)
6184  %xminmax(idsn=enrollment_enr_start_date)
6185  %xminmax(idsn=enrollment_enr_end_date)
6186
6187  *- Order variables -*;
6188  proc sql;
6189       create table dmlocal.&qname as select
6190           datamartid, response_date, query_package, dataset, tag, min, max,
6191             future_dt_n, pre2010_n
6192       from query;
6193  quit;
6194
6195  *- Print query and clear working directory -*;
6196  %clean(savedsn=vital);
6197
6198  ********************************************************************************;
6199  * XTBL_L3_METADATA
6200  ********************************************************************************;
6201  %let qname=xtbl_l3_metadata;
6202
6203  *- Read data set created at top of program -*;
6204  data access;
6205       length sas_base sas_graph sas_stat sas_ets sas_af sas_iml sas_connect
6206              sas_oracle sas_sql sas_mysql sas_postgres sas_teradata sas_odbc $3;
6207       set xtbl_mdata_idsn end=eof;
```

Does anyone else see a "CREATE TABLE" statement and assume that it's a **database** table being created? (Shelley did!)

But in practice, this is creating an output table **within the SAS package**

pcornet

Wrap-up

pcornet

A shortlist of links

- ADAPTABLE base phenotype : https://github.com/ADAPTABLETRIAL/PHENOTYPE
- CDM errata issue tracker: https://github.com/CDMFORUM/CDM-ERRATA/issues
- CDM guidance issue tracker: https://github.com/CDMFORUM/CDM-GUIDANCE/issues
- PCORnet diagnostic query package: https://github.com/PCORnet-DRN-OC/PCORnet-Diagnostic-Query
- PCORnet data characterization query package: https://github.com/PCORnet-DRN-OC/PCORnet-Data-Characterization
- PCORnet Data Committee on GitHub: https://github.com/PCORnet/DataCommittee
- DRN OC home page: https://pcornet.imeetcentral.com/p/aQAAAAAB6T9b