PCORnet CDM Implementation Forum

Thursday, December 3, 2015, 3-4 PM Eastern time

Hosted by Lesley Curtis, PhD

Facilitated by Shelley Rusincovitch and Michelle Smerek

pcornet

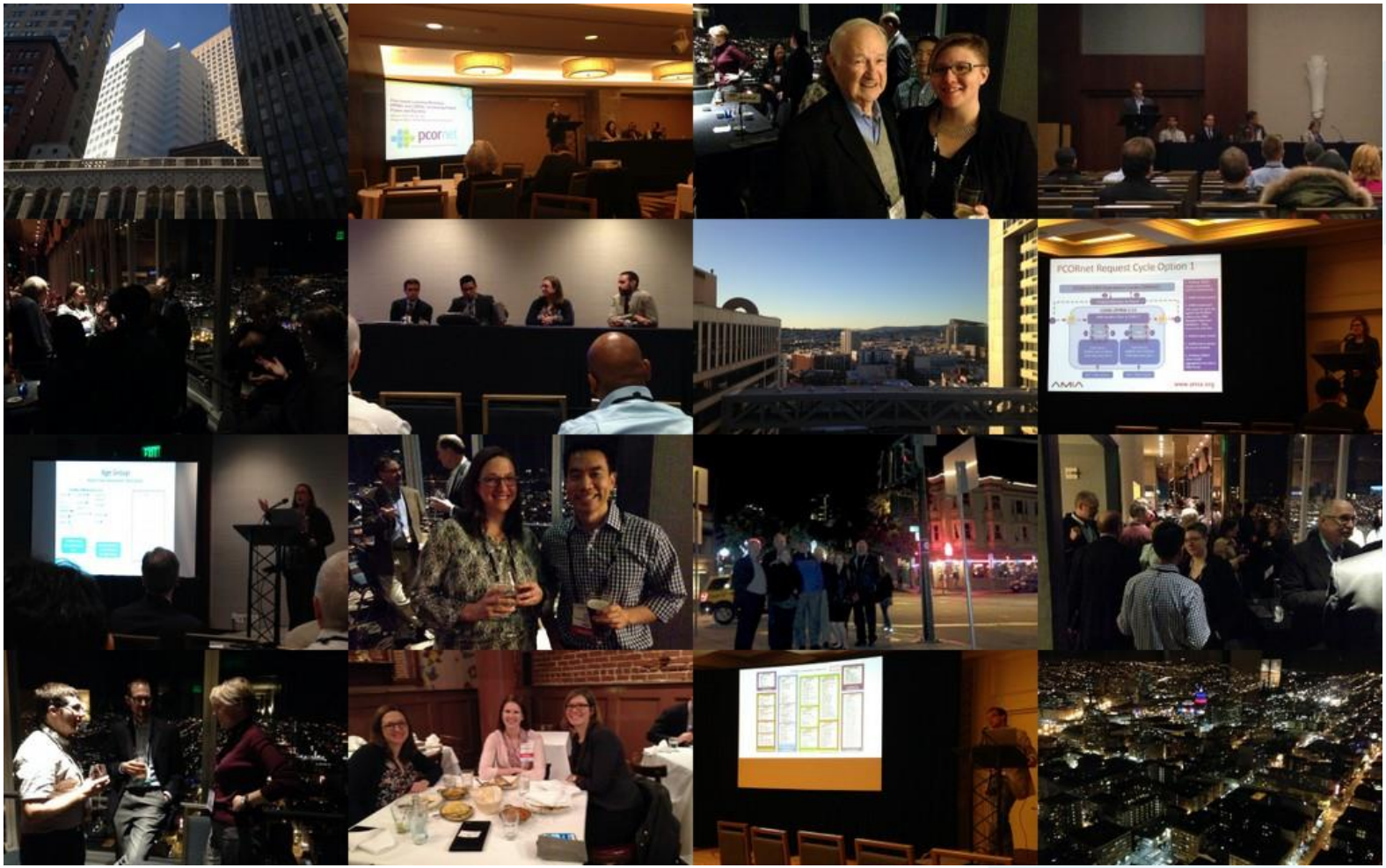
The National Patient-Centered Clinical Research Network

Agenda

- Welcome and announcements
- Index of active CDM forum issues
- Interest group updates
- Smoking and tobacco data elements
- Encounter basis and distinctions between EHR and Claims data sources

pcornet

AMIA Symposium, November 14-18



AMIA photos:
https://pcornet.imeetcentral.com/p/ZgAAAAAAahFs
Publications tracker (including abstracts):
https://pcornet.imeetcentral.com/pcornetmain/dbapp=odap3v5y2j37ru26fg2fylf539212457&ac=h&view=467284

pcornet

Data characterization presentation on December 7

Why it may be of interest to this group:

- Overview of the data characterization package (SAS-based) and process

PCORnet DRN OC-CDRN meeting

Monday, December 7, 11 AM – 12 PM Eastern time

Call-in: 1-650-479-3207 / Access code: 735 866 908

Online:
https://dukemed.webex.com/dukemed/j.php?MTID=maa1bdf8b314534f149efe02cb34319d4

pcornet

ADAPTABLE data strategy on December 11

Why it may be of interest to this group:

- Will include overview of comments from the ADAPTABLE base phenotype specification draft (feedback cycle was November 3-20)

PCORnet ADAPTABLE data strategy discussion

Friday, December 11, 2015, 2:00 PM – 3:00 PM Eastern time

Hosted by Lesley Curtis, PhD, and Schuyler Jones, MD; facilitated by Shelley Rusincovitch and Lisa Eskenazi

Online:
https://dukemed.webex.com/dukemed/j.php?MTID=mfd553360f45df763bbbecdab0429cb2f

Call-in: 1-855-244-8681 / Access code: 731 711 149

pcornet

SAS discussion on December 21

Why it may be of interest to this group:

- Discussion of experiences from pilot data characterization
- Because this program package is run in SAS, discussion will include in-depth experience of SAS architecture/deployment and performance from the CDRNs running the pilots

PCORnet DRN OC-CDRN meeting

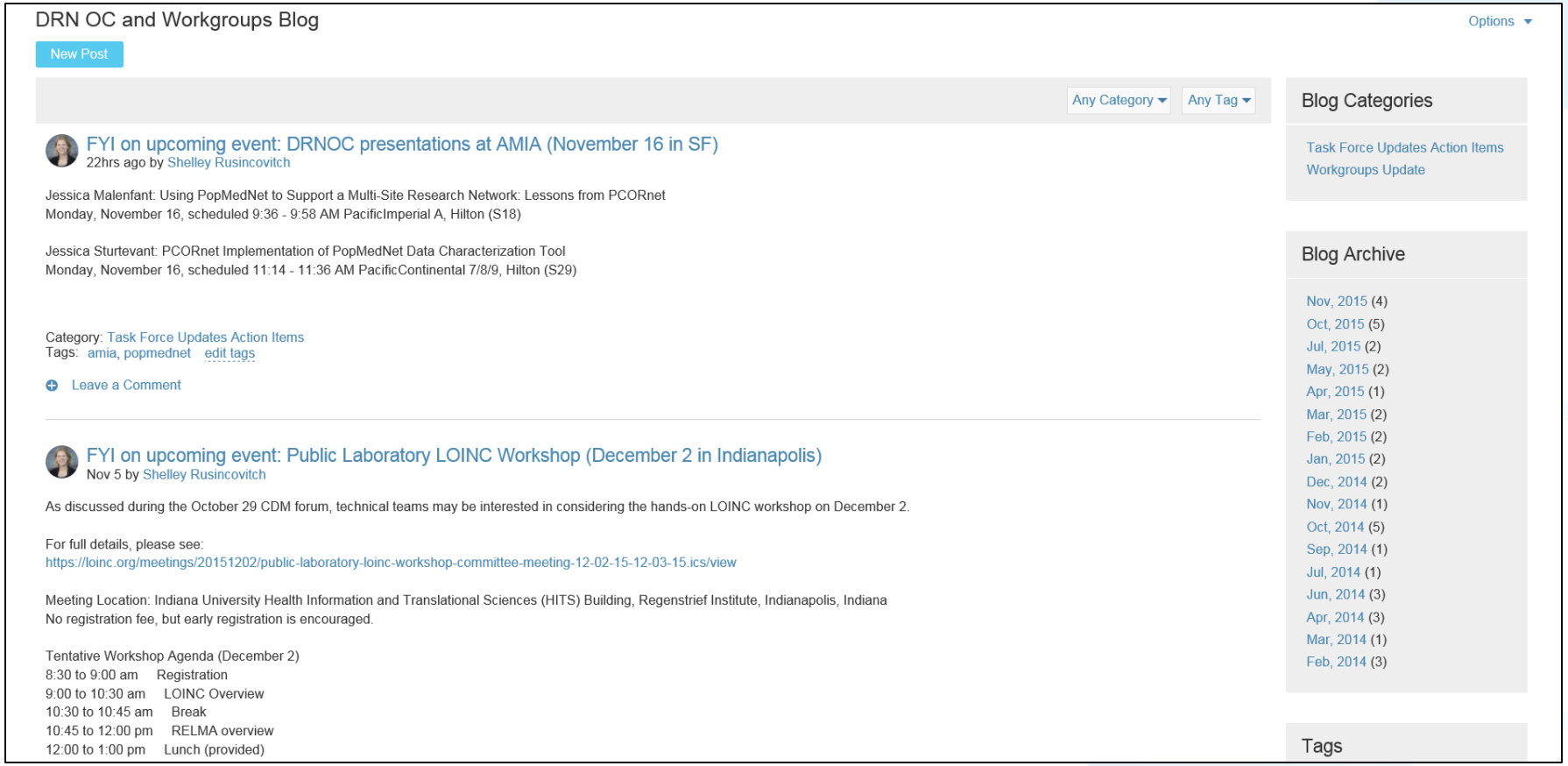
Monday, December 21, 11 AM – 12 PM Eastern time

Call-in: 1-650-479-3207 / Access code: 734 220 460

Online:
https://dukemed.webex.com/dukemed/j.php?MTID=mdc40fe506e9929c39b5431de6331c583

pcornet

DRNOC blog contains updates and links



DRN OC and Workgroups Blog

Options

New Post

Any Category | Any Tag

FYI on upcoming event: DRNOC presentations at AMIA (November 16 in SF)

22hrs ago by Shelley Rusincovitch

Jessica Malenfant: Using PopMedNet to Support a Multi-Site Research Network: Lessons from PCORnet
Monday, November 16, scheduled 9:36 - 9:58 AM PacificImperial A, Hilton (S18)

Jessica Sturtevant: PCORnet Implementation of PopMedNet Data Characterization Tool
Monday, November 16, scheduled 11:14 - 11:36 AM PacificContinental 7/8/9, Hilton (S29)

Category: Task Force Updates Action Items
Tags: amia, popmednet edit tags

Leave a Comment

FYI on upcoming event: Public Laboratory LOINC Workshop (December 2 in Indianapolis)

Nov 5 by Shelley Rusincovitch

As discussed during the October 29 CDM forum, technical teams may be interested in considering the hands-on LOINC workshop on December 2.

For full details, please see:
https://loinc.org/meetings/20151202/public-laboratory-loinc-workshop-committee-meeting-12-02-15-12-03-15.ics/view

Meeting Location: Indiana University Health Information and Translational Sciences (HITS) Building, Regenstrief Institute, Indianapolis, Indiana
No registration fee, but early registration is encouraged.

Tentative Workshop Agenda (December 2)
8:30 to 9:00 am Registration
9:00 to 10:30 am LOINC Overview
10:30 to 10:45 am Break
10:45 to 12:00 pm RELMA overview
12:00 to 1:00 pm Lunch (provided)

Blog Categories

Task Force Updates Action Items
Workgroups Update

Blog Archive

Nov, 2015 (4)
Oct, 2015 (5)
Jul, 2015 (2)
May, 2015 (2)
Apr, 2015 (1)
Mar, 2015 (2)
Feb, 2015 (2)
Jan, 2015 (2)
Dec, 2014 (2)
Nov, 2014 (1)
Oct, 2014 (5)
Sep, 2014 (1)
Jul, 2014 (1)
Jun, 2014 (3)
Apr, 2014 (3)
Mar, 2014 (1)
Feb, 2014 (3)

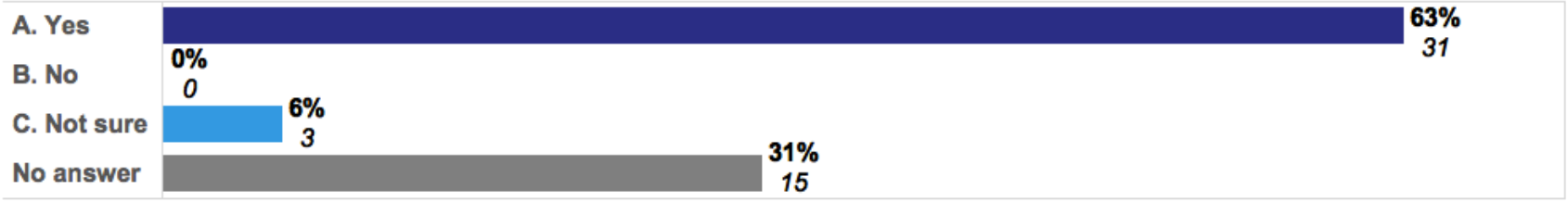
Tags

pcornet

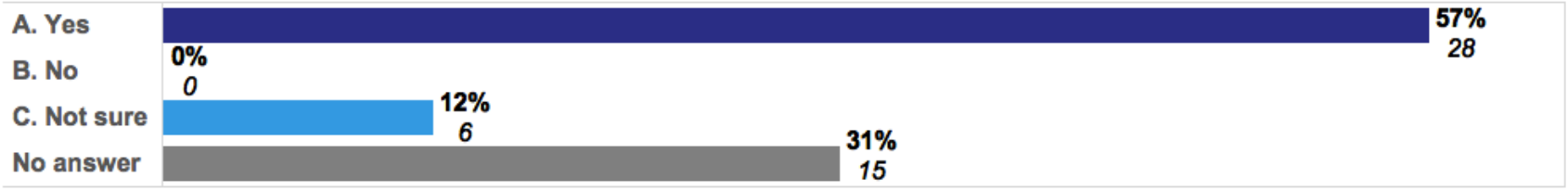
https://pcornet.imeetcentral.com/drnoc-workgroups/blog/

Poll results from CDM Implementation Forum on November 11, 2015 (Code Sharing/Repository)

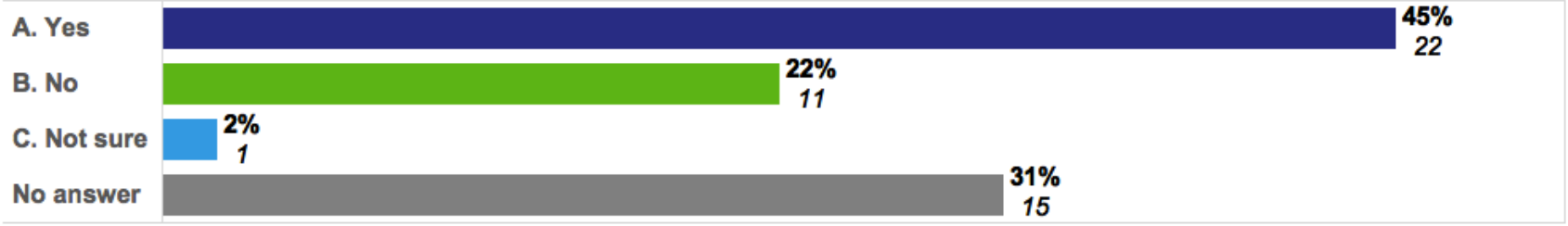
1. Would you (or your team) potentially have interest in USING or IMPROVING UPON PCORnet-related code shared by others? ("Code" might include, but is not limited to, database scripts, processing scripts, client-based utilities, etc.)



2. Would you (or your team) potentially have interest in SHARING PCORnet-related code from your team with others? ("Code" might include, but is not limited to, database scripts, processing scripts, client-based utilities, etc.)



3. Do you (or your team) have experience with a code repository tool (like GitHub or BitBucket)?



CDM Forum Topic Index

pcornet

Outstanding CDM Forum Issues (1 of 2)

Outstanding Issue	Responsible	Date	Actions Taken/Pending
Unclear how to handle auto-generated records (e.g. as in IMO) in CONDITION Table; Unclear whether IMO should be source	Interest Group	TBD	Established interest group (Michelle Smerek facilitating)
Data partners have requested guidance on Medication Mapping conventions	Interest Group	TBD	Established interest group (Michelle Smerek lead)
Data partners have requested guidance on Encounter Classifications	DRNOC and Data Partners	**TODAY** (continuation from Nov 11 forum)	Some overlap with SBAR developed and presented at DRNOC-CDRN meeting on Oct 19
Data partners have requested guidance on death table constraints	DRNOC	TBD	Pending (Shelley Rusincovitch responsible)
Data partners have requested guidance on conventions for representing Smoking and Tobacco History	DRNOC and Data Partners	**TODAY**	Identified legacy data differences from newer MU-mandated structuring

Outstanding CDM Forum Issues (2 of 2)

Outstanding Issue	Responsible	Date	Actions Taken/Pending
Conventions for Datamart Structuring for EHR and Claims Sources need to be defined	DRNOC	**TODAY** (continuation from Nov 11 forum)	SBAR developed and presented at DRNOC-CDRN meeting on Oct 19
SAS implementation expectations need to be defined	DRNOC	Dec 21	Overlap with ADAPTABLE data strategy (session on Oct 30); prior discussion on Nov 11
Data partners are interested in sharing best practices for performance optimization	Forum-facilitated discussion	Dec 21 (SAS basis)	Comment: Best practices tend to be site-specific because different "fingerprints of data" are optimized differently
Data partners have requested guidance on conventions for mapping LOINC to PCORnet common measures	Interest Group	Dec 16	Lab Interest group established
Data partners are interested in sharing best practices for local death data acquisition	Forum-facilitated discussion	TBD	This topic is unrelated to study-specific death data acquisition (such as use of NDI for ADAPTABLE)

CDM Forum Interest Groups

pcornet

Interest Groups List

Active DRNOC-facilitated interest group:

- **Lab Mappings:** Local lab result mappings and LOINC references
- **Med Mappings:** Dispensing and prescribing data, including RxNorm practices, order of preference as brand vs generic

Proposed network-facilitated interest group:

- **CONDITION Table**: Including IMO terminology

pcornet

Interest Groups

- Lab Mapping Activity Survey
  - **34 responses submitted**
  - Interest Group will discuss results and next steps during call on Dec 16th
- Public Laboratory LOINC Workshop
  - Report out
  - Interest Group will discuss how to leverage Workshop information to support network strategies to map local lab results to CDM
- AMIA Lab Mapping Landscape Abstract - status

pcornet

Interest Groups (continued)

- **Medication Mapping**
  - 3 people have expressed interest in participating
  - Exploring use cases that will guide Group activities
  - Email michelle.smerek@duke.edu if interested!

- **CONDITION table**
  - 2 people have expressed interest in participating
  - Email michelle.smerek@duke.edu if interested!

pcornet

Smoking and Tobacco Data Elements

pcornet

Many (most?) source data systems have either smoking or tobacco, **not both**

This field covers Meaningful Use standard:
Any form smoked, but not all tobacco use

These 2 fields covers **any form of tobacco** (smoked and not smoked)

VITAL Table Specification

Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Source*
SMOKING	RDBMS Text(2)	SAS Char(2)	01=Current every day smoker 02=Current some day smoker 03=Former smoker 04=Never smoker 05=Smoker, current status unknown 06=Unknown if ever smoked 07=Heavy tobacco smoker 08=Light tobacco smoker NI=No information UN=Unknown OT=Other	This field is new to v3.0. Indicator for any form of tobacco **that is smoked**. Per Meaningful Use guidance, "…smoking status includes any form of tobacco that is smoked, but not all tobacco use." "'Light smoker' is interpreted to mean less than 10 cigarettes per day, or an equivalent (but less concretely defined) quantity of cigar or pipe smoke. 'Heavy smoker' is interpreted to mean greater than 10 cigarettes per day or an equivalent (but less concretely defined) quantity of cigar or pipe smoke." "…we understand that a "current every day smoker" or "current some day smoker" is an individual who has smoked at least 100 cigarettes during his/her lifetime and still regularly smokes every day or periodically, yet consistently; a "former smoker" would be an individual who has smoked at least 100 cigarettes during his/her lifetime but does not currently smoke; and a "never smoker" would be an individual who has not smoked 100 or more cigarettes during his/her	PCORnet Meaningful Use Core Measures 9 of 13, Stage 1 (2014 definition) http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf [retrieved January 11, 2015]
TOBACCO	RDBMS Text(2)	SAS Char(2)	01=Current user 02=Never 03=Quit/former user 04=Passive or environmental exposure 06=Not asked NI=No information UN=Unknown OT=Other	This field is new to v2.0 with revised value set and field definition in v3.0. Indicator for **any form of tobacco**.	MSCDM with modified field name, field size, and value set
TOBACCO_TYPE	RDBMS Text(2)	SAS Char(2)	01=Smoked tobacco only 02=Non-smoked tobacco only 03=Use of both smoked and non-smoked tobacco products 04=None 05=Use of smoked tobacco but no information about non-smoked tobacco use NI=No information UN=Unknown OT=Other	This field is new to v2.0, with revised value set in v3.0. Type(s) of tobacco used.	MSCDM with modified field size and value set

PCORnet CDM v3.0, pages 39-40 (modified to remove page break).
http://www.pcornet.org/pcornet-common-data-model/

pcornet

Certification Criteria*	
§170.314(a)(11) Smoking status	Enable a user to electronically record, change, and access the smoking status of a patient in accordance with the standard specified at § 170.207(h).

*Additional certification criteria may apply. Review the ONC 2014 Edition EHR Certification Criteria Grid Mapped to Meaningful Use Stage 1 for more information.

Standards Criteria	
§170.207(h)	Coded to one of the following SNOMED CT® codes: (1) Current every day smoker. 449868002 (2) Current some day smoker. 428041000124106 (3) Former smoker. 8517006 (4) Never smoker. 266919005 (5) Smoker, current status unknown. 77176002 (6) Unknown if ever smoked. 266927001 (7) Heavy tobacco smoker. 428071000124103 (8) Light tobacco smoker. 428061000124105

A challenge is that this value set includes multiple concepts (denoted with colors)

pcornet

Meaningful Use Core Measures Measure 9 of 13, Stage 1(2014 Definition). www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf

This slide from 2015-10-08 forum (with thanks again to Daniella Meeker):
https://pcornet.imeetcentral.com/p/ZgAAAAAAaMto

§ 170.315(a)(12) Smoking status p.67

- Laudable to liberalize the use of SNOMED CT codes to represent smoking status beyond the 8 codes used for 2014 edition certification criteria
- Problem that only those 8 codes are permissible for representing smoking status in the Common Clinical Data Set and for electronic transmission in a summary care record.
- The premise that any other smoking status code could be mapped to one of those 8 (as stated in the preamble) is erroneous.
    - Example: For instance, SNOMED 266920004, "trivial cigarette smoker (less than one cigarette/day)" is not a child, in the SNOMED hierarchy, of any of the 8 smoking-related codes required in the 2014 edition certification rule.
- Clarify that this refers to tobacco smoking status, rather than information on the smoking of other substances, since the intent of this criterion appears to be tobacco-specific.
- The Committee recommended a different and shorter (2 questions) measure for tobacco use and exposure than the one previously established as the standard for EHR certification.
- Shorter seems better, even if it requires making a change

Growing recognition of important issues

https://www.healthit.gov/archive/archive_files/HIT%20Standards%20Committee/2015/2015-05-20/HITSC_SSWG_Cert_Rule_2015-05-20_Revised.pdf

pcornet

Tuesday conversation with LPHI/REACHnet

Yes! (with thanks) →

1. **Guidance**: For sites working with CDM v3.0 implementation, are there **best practices/"gotchas"** that would be helpful to share? (especially pertinent to mappings and transformations)

Not at this point →

2. **Maintenance**: Are there **corrections** needed for existing CDM data elements in v3.0?

Yes… →

3. **Assessment**: Is there **uncertainty** about source data practices that would be productive to examine?

Maybe… →

4. **Future**: Are there recommendations for potential future consideration/expansion of the CDM?

pcornet

High-level areas

1. Identifying **concepts** associated with tobacco (eg, mode of tobacco, frequency, etc)

2. **Gaps** between concepts and EHR data sources (includes influence of MU/SNOMED, availability and collection practices within health systems)

3. Question of what are expectations for **tobacco variables in analysis datasets** (such as obesity study) – eg, will analysis variable end up simply as a binary yes/no risk factor?

	MU/SNOMED data collection practices
Present state (current? former?)	Problematic for "heavy tobacco smoker" and "light tobacco smoker"
Frequency (every day? some days?)	Problematic for "heavy tobacco smoker" and "light tobacco smoker"
Duration	Not represented
*Mode of tobacco delivery**	The MU/SNOMED standard is exclusive to smoked tobacco
Exposure (heavy? light?)	Not measured for all responses

*But e-cigarettes are nicotine, not tobacco.

Question about value set for TOBACCO

TOBACCO	RDBMS Text(2)	SAS Char(2)	01=Current user 02=Never 03=Quit/former user 04=Passive or environmental exposure 06=Not asked NI=No information UN=Unknown OT=Other	This field is new to v2.0 with revised value set and field definition in v3.0. Indicator for **any form of tobacco**.	MSCDM with modified field name, field size, and value set

There is no '05' value!

But the "modified" part…

Tobacco	Numeric (1)	1 = Current user 2 = Never 3 = Quit/former user 4 = Passive 5 = Environmental exposure 6 = Not asked 7 = Conflicting	Tobacco status as of the visit date. Unknown values should be left blank. The "Not asked" value should be used only when it is a valid response from your system (e.g. this is a valid value for EPIC). The "Conflicting" value should be used when you receive tobacco information from multiple sources that disagree.

PCORnet CDM v3.0, page 40. http://www.pcornet.org/pcornet-common-data-model/

Mini-Sentinel CDM v4.0, page 43. http://www.mini-sentinel.org/work_products/Data_Activities/Mini-Sentinel_Common-Data-Model.pdf

pcornet

Situation with “smoking” that doesn’t quite fit into MU (case study from 2015-09-24 forum)

What would you do with the value “current smoker” (no mention of frequency)?

You might consider leaving SMOKING = NI

And instead, perhaps TOBACCO = 01 (Current User) and TOBACCO_TYPE = 01 (Smoked Tobacco Only)

VITAL Table Specification

Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Source*
SMOKING	RDBMS Text(2)	SAS Char(2)	01=Current every day smoker 02=Current some day smoker 03=Former smoker 04=Never smoker 05=Smoker, current status unknown 06=Unknown if ever smoked 07=Heavy tobacco smoker 08=Light tobacco smoker NI=No information UN=Unknown OT=Other	This field is new to v3.0. Indicator for any form of tobacco **that is smoked**. Per Meaningful Use guidance, “…smoking status includes any form of tobacco that is smoked, but not all tobacco use.” “‘Light smoker’ is interpreted to mean less than 10 cigarettes per day, or an equivalent (but less concretely defined) quantity of cigar or pipe smoke. ‘Heavy smoker’ is interpreted to mean greater than 10 cigarettes per day or an equivalent (but less concretely defined) quantity of cigar or pipe smoke.” “…we understand that a “current every day smoker” or “current some day smoker” is an individual who has smoked at least 100 cigarettes during his/her lifetime and still regularly smokes every day or periodically, yet consistently; a “former smoker” would be an individual who has smoked at least 100 cigarettes during his/her lifetime but does not currently smoke; and a “never smoker” would be an individual who has not smoked 100 or more cigarettes during his/her	PCORnet Meaningful Use Core Measures 9 of 13, Stage 1 (2014 definition) http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf [retrieved January 11, 2015]
TOBACCO	RDBMS Text(2)	SAS Char(2)	01=Current user 02=Never 03=Quit/former user 04=Passive or environmental exposure 06=Not asked NI=No information UN=Unknown OT=Other	This field is new to v2.0 with revised value set and field definition in v3.0. Indicator for **any form of tobacco**.	MSCDM with modified field name, field size, and value set
TOBACCO_TYPE	RDBMS Text(2)	SAS Char(2)	01=Smoked tobacco only 02=Non-smoked tobacco only 03=Use of both smoked and non-smoked tobacco products 04=None 05=Use of smoked tobacco but no information about non-smoked tobacco use NI=No information UN=Unknown OT=Other	This field is new to v2.0, with revised value set in v3.0. Type(s) of tobacco used.	MSCDM with modified field size and value set

PCORnet CDM v3.0, pages 39-40 (modified to remove page break).
http://www.pcornet.org/pcornet-common-data-model/

pcornet

Encounter basis and distinctions between EHR and Claims data sources

pcornet

Dimensions

- Encounter concept
- Encounter-associated domains (especially DIAGNOSIS and PROCEDURE)
- Encounter field-level classifications for one given data source
- Encounter record structuring when >1 data source available

pcornet

Factors and Possible Approaches to Consider

Network D

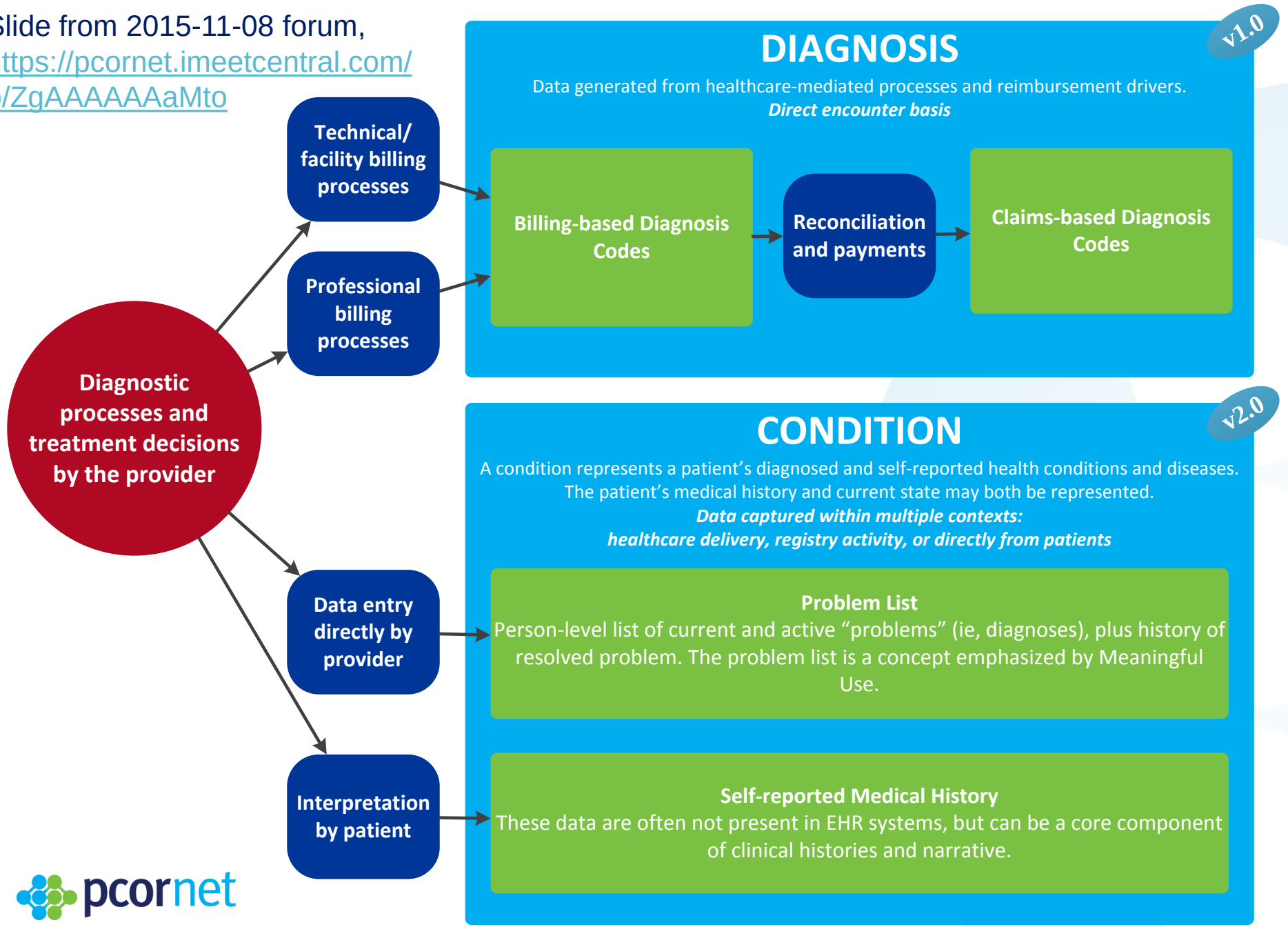
- A discussant stated that the concept of an "encounter" is new to them, and her/his site is integrating hospital and professional billing data.
    - Another discussant noted that the term "encounter" has multiple meanings, and it's necessary to know which one we are talking about in this context.
- The discussant agreed, asking whether the current PCORnet description of "encounter" includes only single encounters with a given provider, or if it would cover all the interactions a patient had associated with one care episode (visit with provider, X-ray, labs, etc.)
    - Another discussant stated that the definition of "provider" is not clear. In this context, are radiologists (for example) providers, or only people who have direct interaction with patients?
- Daniella asked whether introducing the concept of "episode of care" would help in tying related discrete encounters (like those described above) together.
    - The previous discussant noted that one would need to understand how the data associated with such a new concept would be used analytically.
- Shelley noted that it would be important to connect any new development to the current state of the data to put the least amount of burden on the data partners. She noted that Lesley Curtis will be hosting the Forum on December 3rd, providing an opportunity to discuss how these data were used in Mini-Sentinel, and how it's anticipated them may be utilized in PCORnet.

Daniella concluded discussion on this topic by stating that it required much more thought. She suggested that a group be put together to work through the various considerations including: how to interpret transactional data; how to trace provenance, how to attribute encounters to particular providers, etc. Please email shelley.rusincovitch@duke.edu if you are interested in participating in the Encounter Interest Group.

pcornet

From summary by Michelle Smerek of 2015-11-11 Forum, https://pcornet.imeetcentral.com/p/aQAAAAACkQqv

Slide from 2015-11-08 forum, https://pcornet.imeetcentral.com/p/ZgAAAAAAaMto



pcornet

Field-level classification decisions…

…and why this is tricky for claims.

Discussion point: What are the factors that must be considered when deciding whether:

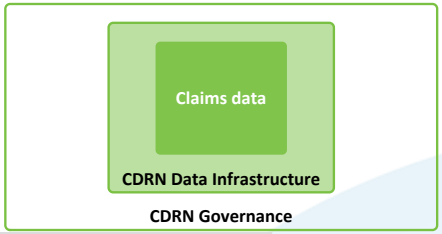
To create one “reconciled” inpatient encounter record

vs.

Mapping the hospital claims to IP and the provider claims to OT, in an “unreconciled” fashion?

pcornet

SBAR: Background (continued)



- Both EHR data and claims data have the concept of “encounters” (interactions between patients and providers within the context of healthcare delivery)
- A patient could have encounter data in the EHR and claims, including associated diagnosis, procedure codes, etc
- Encounter data from each source may be **duplicative**
  - eg, hospitalization from 2/1/2015-2/5/2015 is in EHR; the same hospitalization is in the claims data
- Encounter data may be in **conflict**
  - eg, EHR data has a discharge date of 2/5/2015, claims data says 2/6/2016.

This slide from 2015-10-19 DRNOC-CDRN meeting:
https://pcornet.imeetcentral.com/p/ZgAAAAAAaAfc

pcornet

Option 2: Single datamart **without reconciliation of encounters**

- Full duplication on every table where both claims and EHR data are available.
- This solution would likely involve both record-level flags for source provenance, plus metadata about duplication present in datamart (potentially extend HARVEST table)
- Pros:
  - Less burden upon data partner for reconciliation; likely the preferred option
- Cons:
  - Some CDRNs will not have permission to comingle claims data in the foundational datamart
  - Duplication may impact analyses related to encounter (healthcare utilization) and procedure data
    - Note: Some EHR-only data sources do not reconcile encounter basis for facility vs. professional billing data streams; therefore, the issue of duplication is likely to be widely present

pcornet

This slide from 2015-10-19 DRNOC-CDRN meeting:
https://pcornet.imeetcentral.com/p/ZgAAAAAAaAfc

Option 3. Single datamart with **complete integration of encounter basis**

- The CDRN has transformed and reconciled the data so there is no duplication between claims and billing on any table
- Pro: May not require additional modification of analytic tools
- Cons:
  - Involves significant burden for the data partner
  - Given the complexity of reconciliation, it is possible for data partner to implement with poor quality
  - Some CDRNs will not have permission to comingle claims data in the foundational datamart

pcornet

This slide from 2015-10-19 DRNOC-CDRN meeting:
https://pcornet.imeetcentral.com/p/ZgAAAAAAaAfc

Coming back to the framing…

Probably so! →

1. **Guidance**: For sites working with CDM v3.0 implementation, are there **best practices/"gotchas"** that would be helpful to share? (especially pertinent to mappings and transformations)

Not at this point →

2. **Maintenance**: Are there **corrections** needed for existing CDM data elements in v3.0?

Yes… →

3. **Assessment**: Is there **uncertainty** about source data practices that would be productive to examine?

Maybe… →

4. **Future**: Are there recommendations for potential future consideration/expansion of the CDM?

pcornet

Next CDM Forum

Thursday, January 7, 2016, 1–2 PM Eastern

Hosted by Keith Marsolo, PhD; facilitated by Shelley Rusincovitch and Michelle Smerek

ONLINE:
https://dukemed.webex.com/dukemed/j.php?MTID=m5afcc4c950962b9a87788dcffde5beae

PHONE: 1-855-244-8681 / Access code: 730 227 049

pcornet