Michael E. Matheny, MD, MS, MPH

Associate Director for Analytics, VINCI
Associate Director, Advanced Fellowship in Medical Informatics
Tennessee Valley Healthcare System VA

Director, Center for Population Health Informatics
Associate Professor, Departments of Biomedical Informatics, Medicine, and Biostatistics
Vanderbilt University

Twitter: @MichaelEmatheny
Email: michael.Matheny@va.gov, michael.Matheny@vumc.org

Identifying Acute Inpatient Stays within the CDW Inpatient Domain & Inpatient OMOP Visit Occurrence Transformation

Questions

- How much have you used the VA OMOP Data Resource?
  - None
  - A Little
  - Some
  - A Majority
  - All the time!

- Do you (or your team) have a code transformation to use inpatient data in the VA written for current or prior projects?
  - Yes
  - No

- Provide background on OMOP

- Provide an overview of the ViSTa/CPRS/CDW Inpatient Domain
  - Describe Key Challenges in the use of this data
  - Describe Research & Operations Recommendations

- Provide an Overview of the CDW Inpatient -> OMOP VISIT_OCCURRENCE Transformation

Objectives

OMOP Background

- Provides a common framework for groups working within a data source to share 'data transformation infrastructure'

- Abstracts a portion of the need for detailed source data knowledge, data transformation, and data quality assessments away from source data representations

- Allows externally developed statistical analyses and analytic tools to be used and deployed within a health care system

Rationale for a Common Data Model

- Broadest Coverage & Least Data Fidelity Loss for Comparative Effectiveness

- Robust Open Source Development Community
  - Evolving the common data model
  - Developing analytic and visualization tools for cohort creation, data visualization, and large-scale analytics (OHDSI)

- Data transformation, content mapping, and quality assurance activities are applicable to any future CDM or standard

Ogunyemi OI, et al. Med Care. 2013 Aug;51(8 Suppl 3):S45-52
Huser V, et al. AMIA Annu Symp Proc 2013;648-56.
Garza, et al. J. Biomed. Inform. 2016;64:333-341
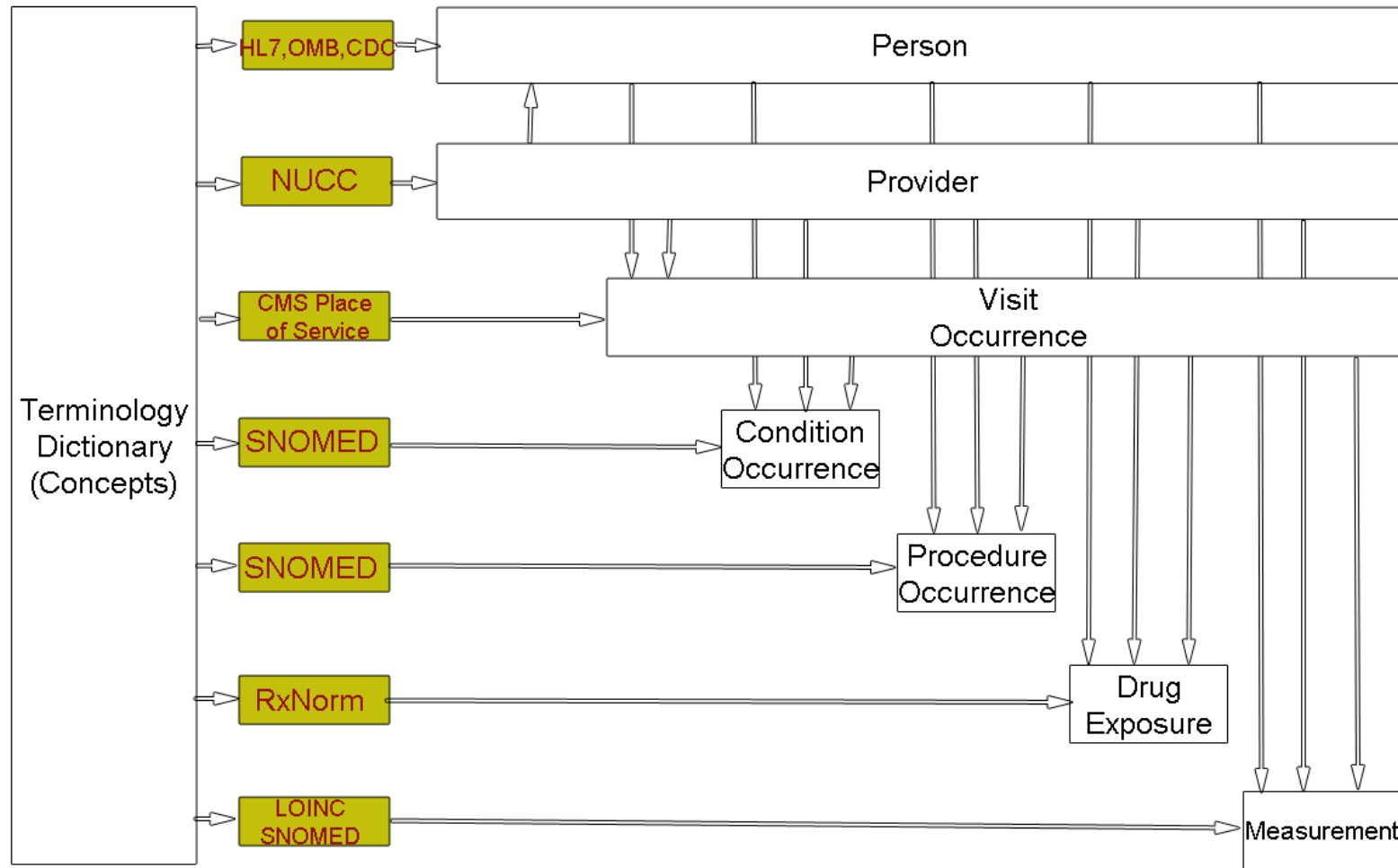Voss EA, et al. J Am Med Inform Assoc. 2015 May;22(3):553-64
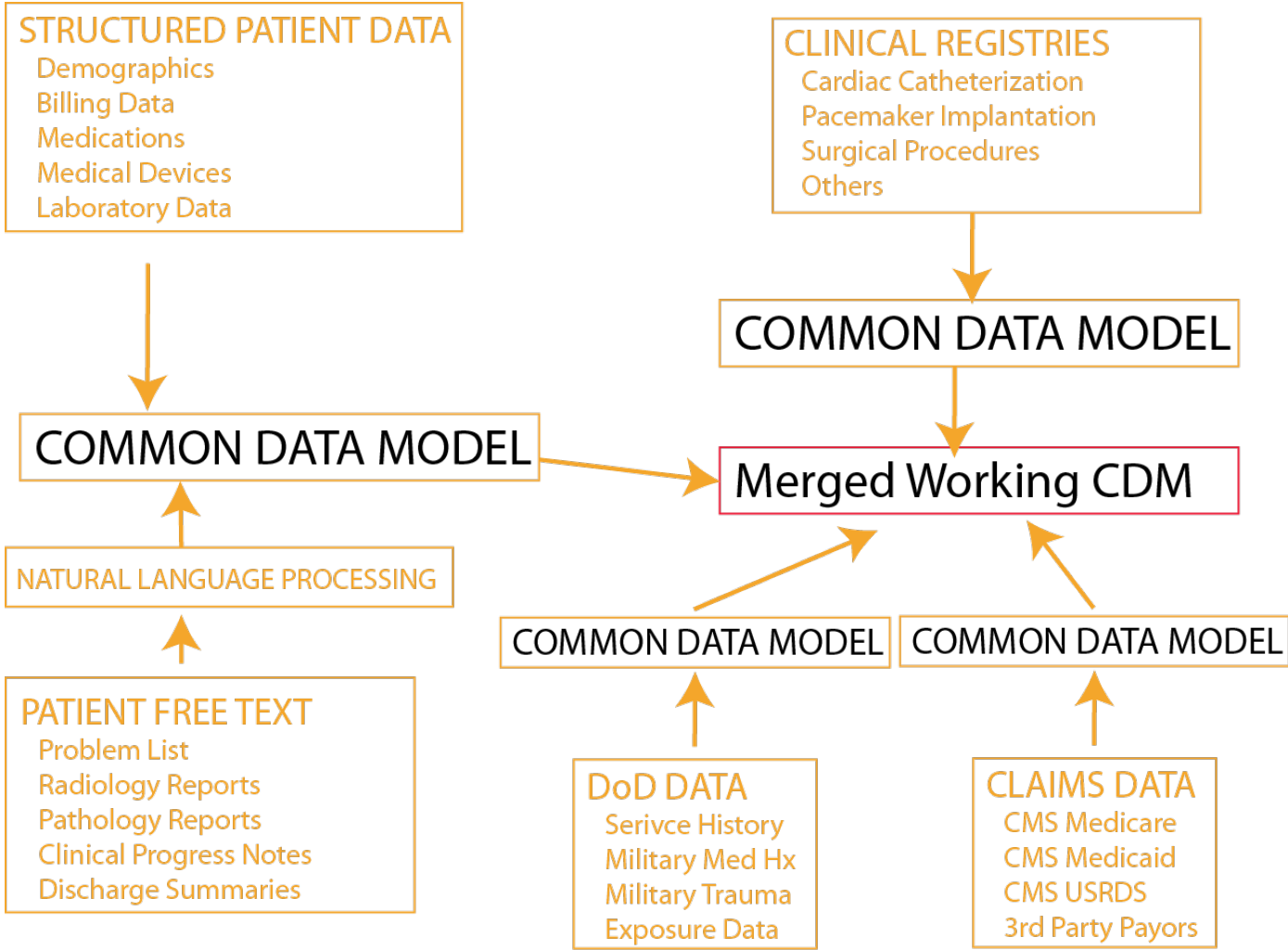
# Rationale for OMOP
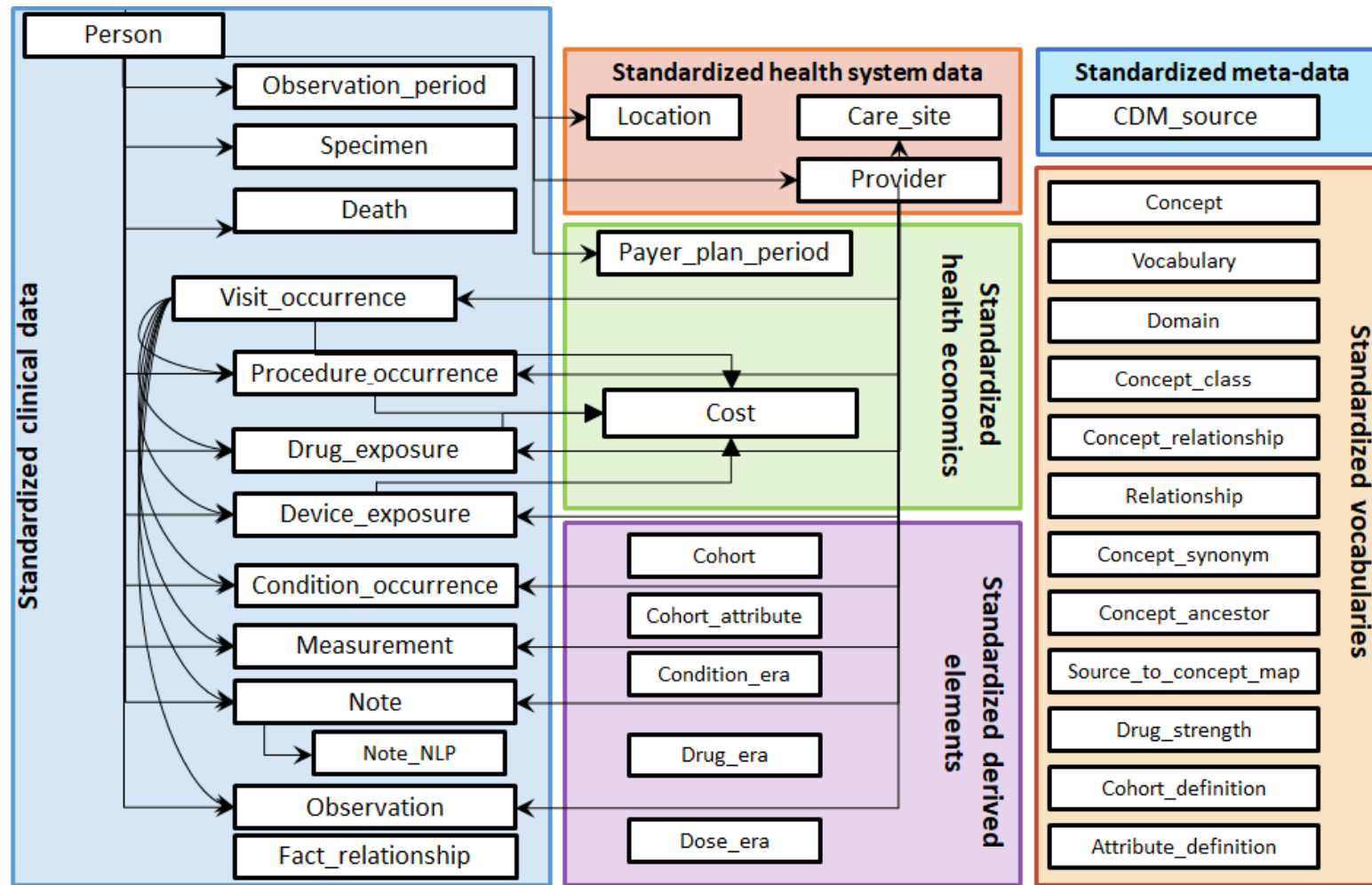
OMOP CDM Version 5.3 Schema

OMOP Data Relationships

Acute Inpatient Stays

- A very common frame of analysis for observational cohort analyses, because of increased granularity of data within this data window

- High volume of research and operational use

- Many Research Extracts and Operational Data Cubes for Inpatient Stays

Acute Inpatient Encounter

Count

* 2018 – partial year query done 11/7/18
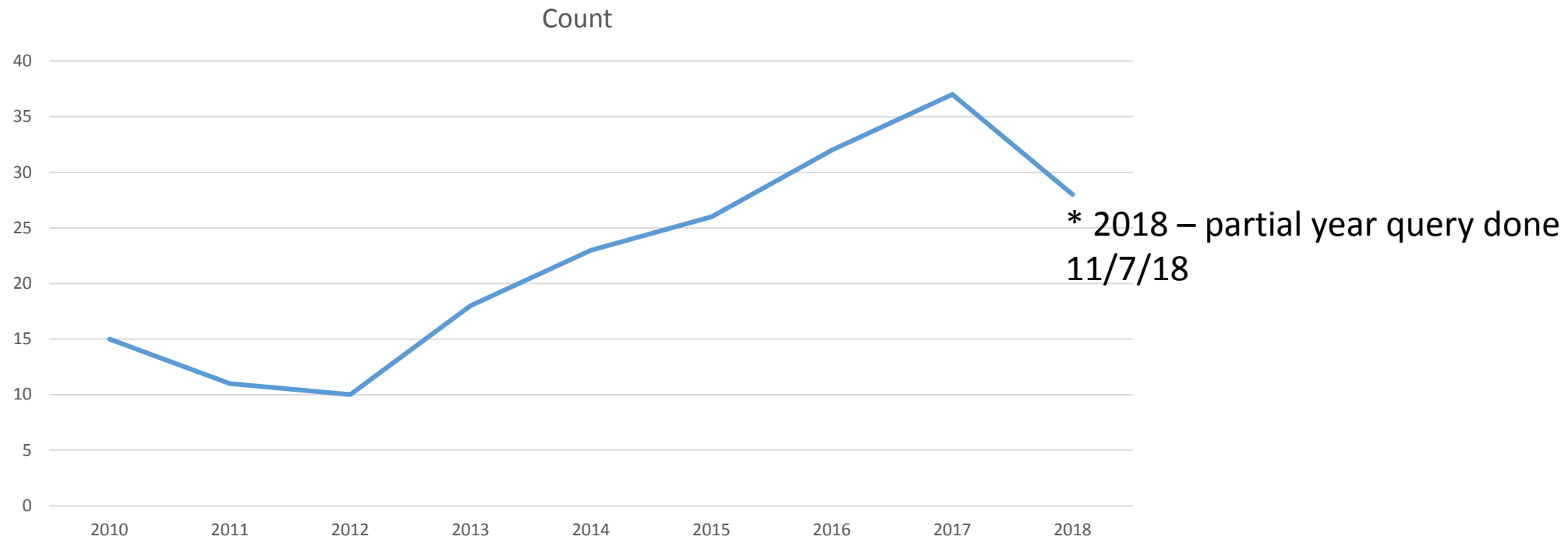
Pubmed Search Criteria: veteran[Title/Abstract] AND (inpatient[Title/Abstract] OR hospitalized[Title/Abstract]) AND ("2018/01/01"[Date - Publication] : "2019/01/01"[Date - Publication])

Inpatient VA Publications

CDW Inpatient 3.0 Domain

- CDW's primary mandate is to extract the underlying source production health care data to make it available to operational and research users

- Although all details are beyond scope of this presentation, the CDW team has gone through 3 iterations of the Inpatient domain and already handled **tremendous** source data complexity before providing the domain in CDW (Richard Pham, Trinity Hall, Steven Anderson)

- Conceptually, data warehousing is **<u>DIFFERENT</u>** from common data model repositories, such as OMOP, I2B2, or PCORNet. There are different trade-offs and requirements for source data and target data representations, usability for which use cases, and considerations for handling and merging massive data volumes.

CDW – Inpatient 3.0
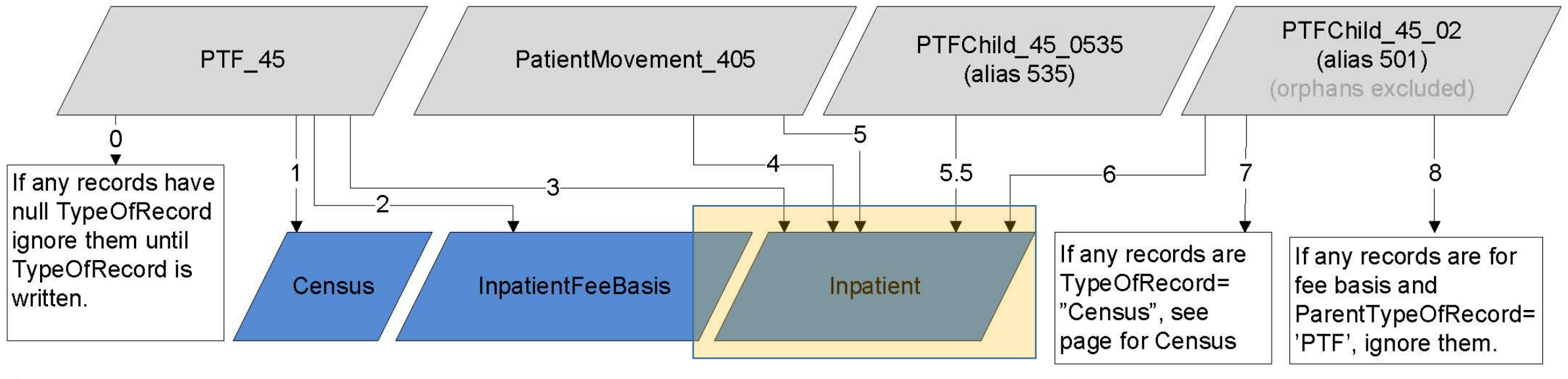
CDW – Inpatient 3.0

- Inpat.Inpatient was calculated from admission and discharge movements from ALL possible
  - 501 Records (File 45.02)
  - 535 Records (File 45.0535)
- Because File 45 has different answers depending on what is used, CDW **<u>provided all versions</u>** to allow maximum flexibility for end users:
  - ProvisionalMovement
  - Inpatient501Transaction
  - SpecialtyTransfer
  - Inpatient535Transaction
  - PatientTransfer
- CDW Recommendations:
  - Use Inpatient (45) Specialty for Discharge
  - Use File 45.02 Specialty for Admission
  - Use File 45.0535 for Ward Locations (Admit & DC)
- For additional details, see the CDW documentation at https://vaww.cdw.va.gov/metadata/

CDW Inpatient 3.0 – Design Considerations

OMOP Design Choice – Focus on:

- Because we don't currently have to address present tense assignments (nightly updates with real-time data), we can use:

- Inpat_Inpatient: Summary level of the inpatient stay, includes admission & discharge dates and other key data

- Inpat_PatientTransfer: One of the source of truths for a row per change in location or type of care for a patient while they are in inpatient status.

CDW – Inpatient 3.0 -> OMOP

Inpatient

[InpatientSID]
,[Sta3n]
,[PatientSID]
,[AdmitDateTime]
,[DischargeDateTime]

,[AdmitSourceSID]
,[TransferFromFacility]
,[AdmitMASMovementTypeSID]
,[AdmitFacilityMovementTypeSID]
,[AdmitFromInstitutionSID]
,[AdmitRoomBedSID]

,[DischargeStatus]
,[DispositionType]
,[DischargeFromFacility]
,[PlaceOfDispositionSID]
,[TransferToFacility]
,[AdmitWardLocationSID]

,[DischargeMASMovementTypeSID]
,[DischargeFacilityMovementTypeSID]
,[DischargeToInstitutionSID]
,[DischargeAttendingPhysicianSID]
,[DischargeFromSpecialtySID]
,[DischargeFromService]
,[ProviderSID]
,[DischargeWardLocationSID]
,[DischargeSpecialtySID]

,[PrincipalDiagnosisICD9SID]
,[PrincipalDiagnosisICD10SID]

Patient Transfer

[PatientTransferSID]
,[Sta3n]
,[InpatientSID]
,[OrdinalNumber]
,[PatientSID]
,[PatientTransferDateTime]

,[MASMovementTypeSID]
,[FacilityMovementTypeSID]

,[TransferInstitutionSID]
,[LosingSpecialtySID]
,[LosingWardLocationSID]
,[LosingWardCDR]

,[GainingWardLocationSID]
,[RoomBedSID]
,[DischargeMovementFlag]

,[PrimaryPhysicianStaffSID]
,[AttendingPhysicianStaffSID]

Summary level of the inpatient stay, includes admission & discharge dates and other key data

A row per change in location or type of care for a patient while they are in inpatient status.

CDW Source (Inpatient Domain)

Inpatient

Patient Transfer

[InpatientSID]
,[Sta3n]
,[PatientSID]
,[AdmitDateTime]
,[DischargeDateTime]

,[AdmitSourceSID]
,[TransferFromFacility]
,[AdmitMASMovementTypeSID]
,[AdmitFacilityMovementTypeSID]
,[AdmitFromInstitutionSID]
,[AdmitRoomBedSID]

,[DischargeStatus]
,[DispositionType]
,[DischargeFromFacility]
,[PlaceOfDispositionSID]
,[TransferToFacility]
,[AdmitWardLocationSID]

,[DischargeMASMovementTypeSID]
,[DischargeFacilityMovementTypeSID]
,[DischargeToInstitutionSID]
,[DischargeAttendingPhysicianSID]
,[DischargeFromSpecialtySID]
,[DischargeFromService]
,[ProviderSID]
,[DischargeWardLocationSID]
,[DischargeSpecialtySID]

,[PrincipalDiagnosisICD9SID]
,[PrincipalDiagnosisICD10SID]

[PatientTransferSID]
,[Sta3n]
,[InpatientSID]
,[OrdinalNumber]
,[PatientSID]
,[PatientTransferDateTime]

,[MASMovementTypeSID]
,[FacilityMovementTypeSID]

,[TransferInstitutionSID]
,[LosingSpecialtySID]
,[LosingWardLocationSID]
,[LosingWardCDR]

,[GainingWardLocationSID]
,[RoomBedSID]
,[DischargeMovementFlag]

,[PrimaryPhysicianStaffSID]
,[AttendingPhysicianStaffSID]

Summary level of the inpatient stay, includes admission & discharge dates and other key data

A row per change in location or type of care for a patient while they are in inpatient status.

CDW Source (Inpatient Domain)

Inpatient

Patient Transfer

[InpatientSID]
,[Sta3n]
,[PatientSID]
,[AdmitDateTime]
,[DischargeDateTime]

,[AdmitSourceSID]
,[TransferFromFacility]
,[AdmitMASMovementTypeSID]
,[AdmitFacilityMovementTypeSID]
,[AdmitFromInstitutionSID]
,[AdmitRoomBedSID]

,[DischargeStatus]
,[DispositionType]
,[DischargeFromFacility]
,[PlaceOfDispositionSID]
,[TransferToFacility]
,[AdmitWardLocationSID]

,[DischargeMASMovementTypeSID]
,[DischargeFacilityMovementTypeSID]
,[DischargeToInstitutionSID]
,[DischargeAttendingPhysicianSID]
,[DischargeFromSpecialtySID]
,[DischargeFromService]
,[ProviderSID]
,[DischargeWardLocationSID]
,[DischargeSpecialtySID]

,[PrincipalDiagnosisICD9SID]
,[PrincipalDiagnosisICD10SID]

[PatientTransferSID]
,[Sta3n]
,[InpatientSID]
,[OrdinalNumber]
,[PatientSID]
,[PatientTransferDateTime]

,[MASMovementTypeSID]
,[FacilityMovementTypeSID]

,[TransferInstitutionSID]
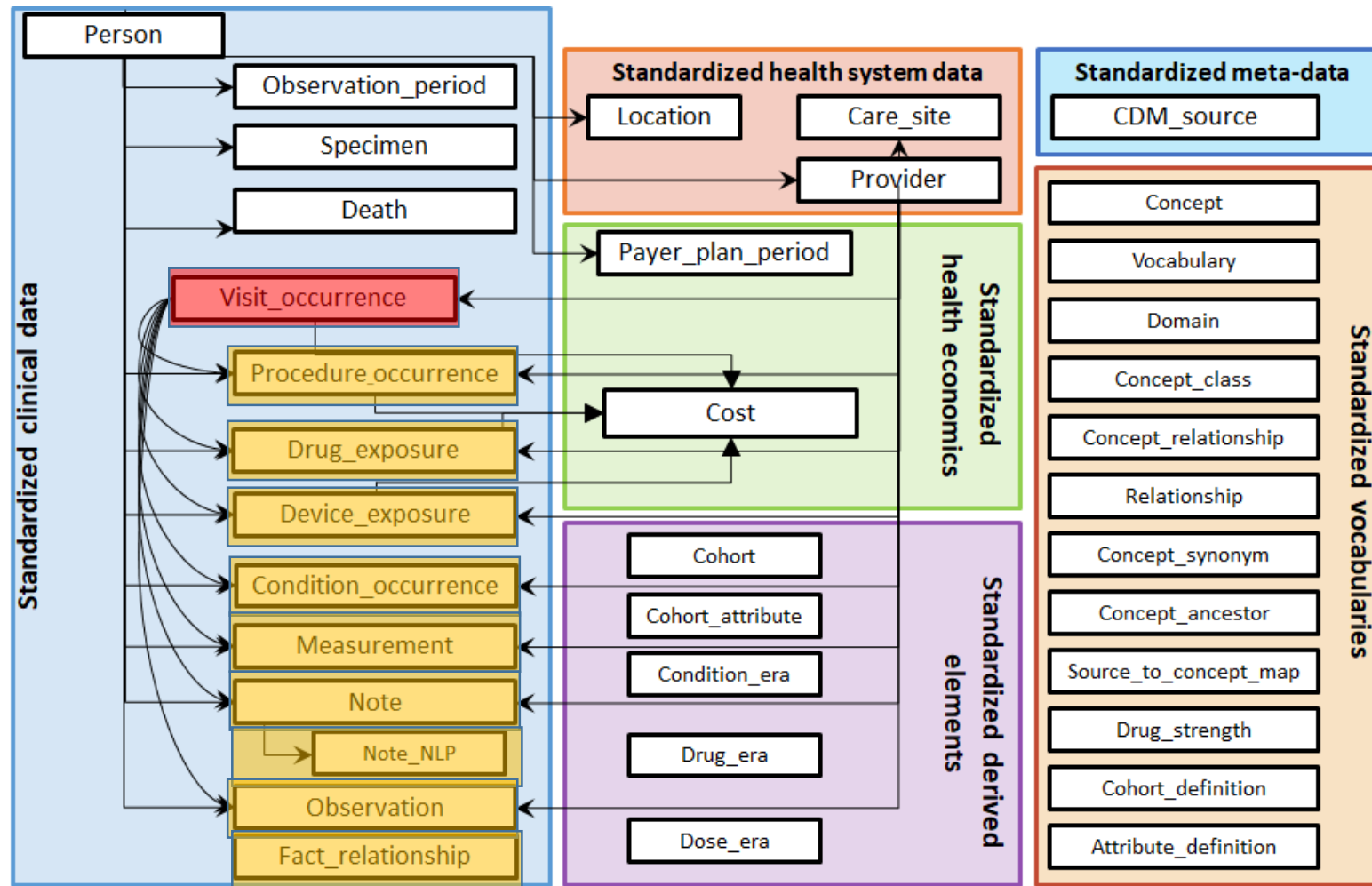,[LosingSpecialtySID]
,[LosingWardLocationSID]
,[LosingWardCDR]

,[GainingWardLocationSID]
,[RoomBedSID]
,[DischargeMovementFlag]

,[PrimaryPhysicianStaffSID]
,[AttendingPhysicianStaffSID]

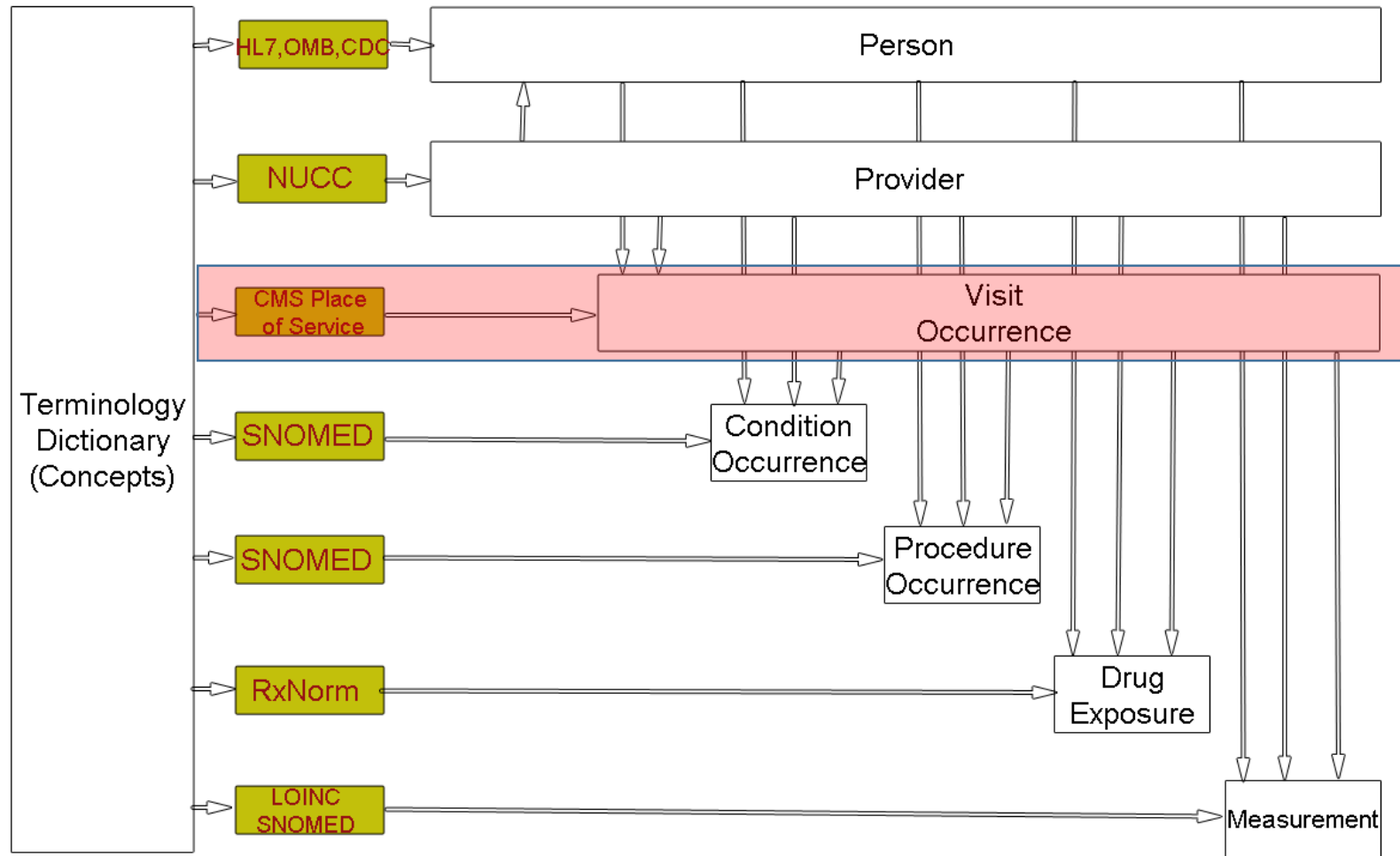**Summary level of the inpatient stay, includes admission & discharge dates and other key data**

A row per change in location or type of care for a patient while they are in inpatient status.

CDW Source (Inpatient Domain)

Inpatient

[InpatientSID]
,[Sta3n]
,[PatientSID]
,[AdmitDateTime]
,[DischargeDateTime]

,[AdmitSourceSID]
,[TransferFromFacility]
,[AdmitMASMovementTypeSID]
,[AdmitFacilityMovementTypeSID]
,[AdmitFromInstitutionSID]
,[AdmitRoomBedSID]

,[DischargeStatus]
,[DispositionType]
,[DischargeFromFacility]
,[PlaceOfDispositionSID]
,[TransferToFacility]
,[AdmitWardLocationSID]

,[DischargeMASMovementTypeSID]
,[DischargeFacilityMovementTypeSID]
,[DischargeToInstitutionSID]
,[DischargeAttendingPhysicianSID]
,[DischargeFromSpecialtySID]
,[DischargeFromService]
,[ProviderSID]
,[DischargeWardLocationSID]
,[DischargeSpecialtySID]

,[PrincipalDiagnosisICD9SID]
,[PrincipalDiagnosisICD10SID]

Patient Transfer

[PatientTransferSID]
,[Sta3n]
,[InpatientSID]
,[OrdinalNumber]
,[PatientSID]
,[PatientTransferDateTime]

,[MASMovementTypeSID]
,[FacilityMovementTypeSID]

,[TransferInstitutionSID]
,[LosingSpecialtySID]
,[LosingWardLocationSID]
,[LosingWardCDR]

,[GainingWardLocationSID]
,[RoomBedSID]
,[DischargeMovementFlag]

,[PrimaryPhysicianStaffSID]
,[AttendingPhysicianStaffSID]

Summary level of the inpatient stay, includes admission & discharge dates and other key data

A row per change in location or type of care for a patient while they are in inpatient status.

CDW Source (Inpatient Domain)

OMOP Transformation

OMOP CDM Version 5.3 Schema

Source: OMOP Common Data Model Draft Specification Version 5.3, November 2017

Visit Occurrence Standard Vocabulary

OMOP Visit Occurrence - Overview

OMOP Visit Occurrence – Field Descriptions

- Pharmacy
- Telehealth
- Homeless Shelter
- Office
- Home
- Emergency Room
- Ambulatory Surgical Center
- Hospice
- Community MHC
- Outpatient Rehab Facility
- Urgent Care Facility

- Inpatient Hospital
- Inpatient Psych Facility

- Assisted Living Facility
- Group Home
- Skilled Nursing Facility
- Nursing Facility
- Residential Psych Treatment Center
- Custodial Care Facility (Domicilliary)

CMS Place Of Service (Partial Excerpt)

- How to determine what type of encounter the patient was experiencing
  - Separation of Acute Inpatient & Long Term Care is critical

- How to handle "attached" data? (administrative codes, laboratory tests, etc.)

OMOP Conceptual Challenges

- Remove exact duplicate records
- Remove records where PatientSID does not match between Inpatient and Patient Transfer
- Generate 2 rows of data from each Inpatient record
  - a 'transfer' from outside (source varies) to inpatient
  - a 'transfer' (discharge) from inpatient to outside (source varies).
- Merge records between Inpatient and Patient Transfer

Data Cleaning – Stage 1

Inpatient <-> Patient Transfer

- Determination of type of care being transferred from and to ('MovementType', Indeterminate, ASIH, Ward, etc)
- Determination of Service being transferred from and to ('Domiciliary', 'Medicine','NHCU',etc)
- Developed a set of Boolean rules for movement type and service equivalence (collapsibility)

Data Type Harmonization – Stage 2

- Sequence all data for each patientSID
  - inpatient focus so can use site specificity to advantage
- Merge all sequences same place of service type
- Merge acute stay hospitalization sub-types together
- Merge long-term care Nursing Home type together
- Merge Domiciliary type stays together

Encounter Type Merger – Stage 3

- Analyze all overlapping records and attempt to disambiguate
- Longer encompassing records are more likely to be non-acute care

  - if enough characteristics are non-acute consistent for outer & acute for inner then break encounters with inner record to be acute and outer split on both sides.

  - For acute overlapping, merge

  - For non-acute overlapping, merge

Handling of Overlapping Encounters – Stage 4

- Source Data (10/01/1999+):   13,644,795
- Transformed Data: 15,021,440

- Transformation Magnitude
  - Admit & Discharge Date/Time Stamps Unchanged: 3,385,696
  - Admit & Discharge Date Stamps Unchanged: 12,911,799
  - Source Records
    - Merged (2+ to 1):   794,999
    - Split (1 to 2+):   811,301

Summary Transformation Characteristics

- All administrative and other data that are attached to encounter but that exist in other OMOP tables retain their original source date/time stamp (source fidelity decision).

- They are attached to the original encounter segment from the source data (maintain source linkage)

- When the resulting encounter is not a one-to-one mapping to associated data, first match to encounter date window, if no match, leave associated data without mapped VISIT pointer

"Attached" Data Handling

- Inpatient data from source are not usable without significant cleaning

- Initial QA efforts were conducted but additional validation efforts ongoing and welcome collaboration

- Substantially disambiguates inpatient encounter records

- Very complex logic, over 4000 lines of SQL code

Summary

- Ease of Use – 1-3 database joins for most applications to start using inpatient data in OMOP

- Efficient – heavy processing load once for transform allows faster downstream queries. Also frees up space for other intermediate table products in research databases

- Reproducible – studies using this transform can be assured of direct comparability to other studies derived from this source and for consistent implementation

- Transparency – documentation and transformation logic are available to VA users

- Enforced Convention – We elected to transform the data using one of the possible paths the source data allows, but there are different ground truths for different use cases, so it may not be applicable to your use case

- Consensus Building - We welcome the field to provide feedback and criticism of the transform… no transform can satisfy all use cases, but we seek to support the majority of research use cases

Big Picture
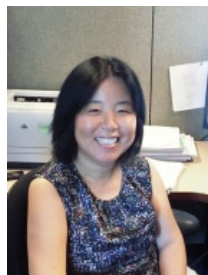
Aize
Cao

Scott
DuVall

Fern
FitzHenry

Kristine
Lynch

Michael
Matheny

Sharidan
Parr

Jason
Denton

Liz
Hanchrow

Kushan
Hewa

Abby
Hillard

Steven
Johnson

Vincent
Messina

Daniel
Park

Ben
Viernes

VINCI OMOP Team