



Identification of Outpatient Workload in CDW (Corporate Data Warehouse)

**Data Quality Analysis Team
VHA OIA HIG Data Quality
Business Product Management**

Laureen Kennedy
Trinity Hall

May 16, 2013

Poll Question

- What best describes your position at the VA?
 - Researcher
 - Clinician
 - Administrator/policymaker
 - Other

Questions Addressed by Session

- Why do we need a CDW Outpatient Workload Query?
- What does the CDW Outpatient Workload Query do?
 - Criteria for workload in VistA
 - Translation to a Relational Database
- What did we learn developing & validating our query?
- How can you use the Query?
- Conclusions

Questions Addressed by Session

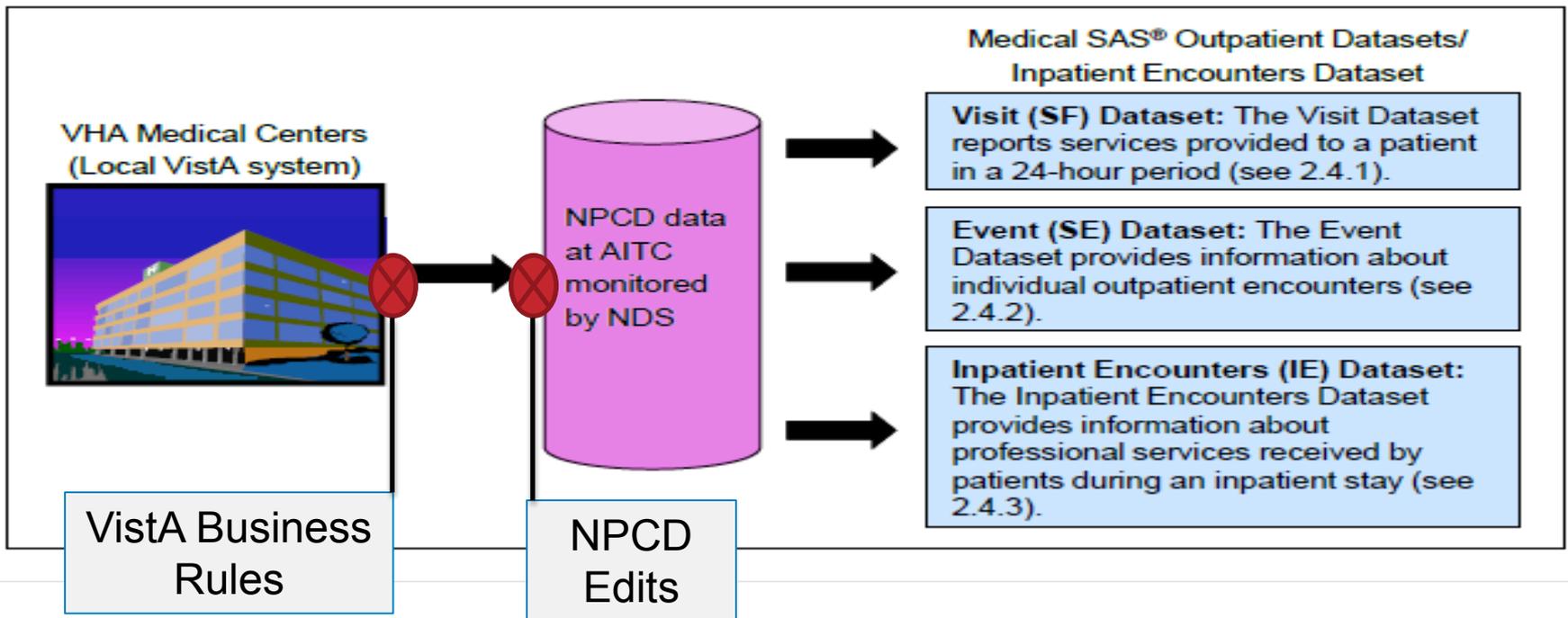
- Why do we need a CDW Outpatient Workload Query?
- What does the CDW Outpatient Workload Query do?
 - Criteria for workload in VistA
 - Translation to a Relational Database
- What did we learn developing & validating our query?
- How can you use the Query?
- Conclusions

Why do we need a CDW Outpatient Workload Query?

- Outpatient MedSAS Datasets contain VHA outpatient workload.
 - Derived from VistA outpatient data which is transmitted to the National Patient Care Database (NPCD).
- NPCD is being decommissioned in 2014. CDW contains VistA outpatient data, but in a different format.

Why do we need a CDW Outpatient Workload Query?

Figure 1. Data progression from VHA Medical Centers to the AITC and to the Medical SAS® Outpatient Datasets



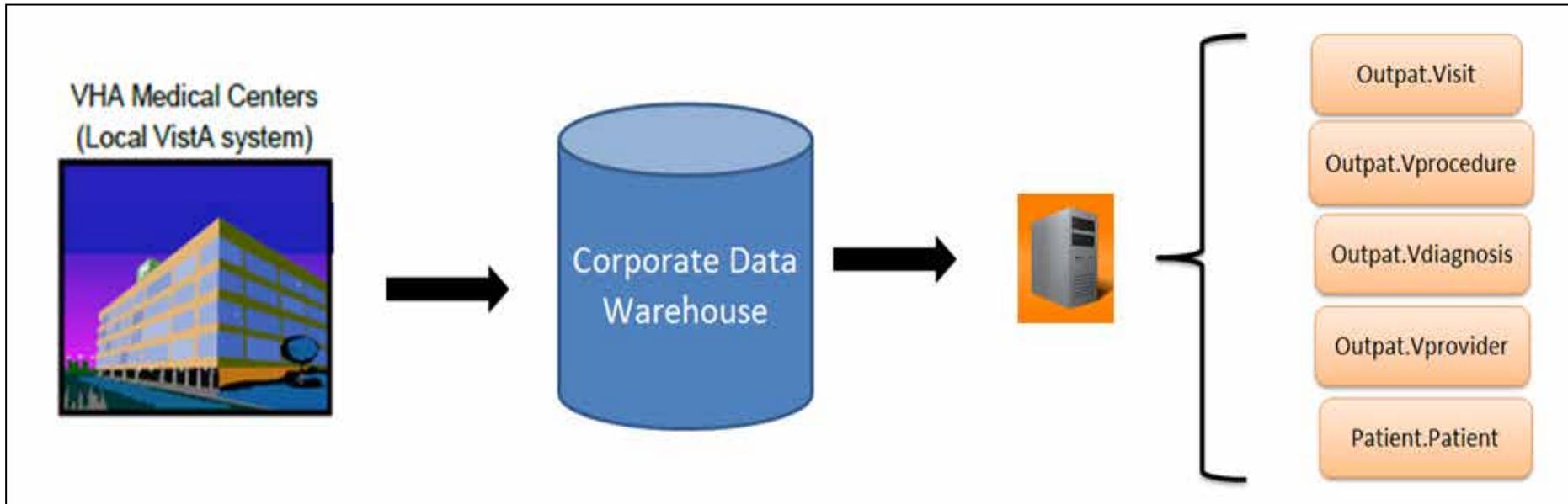
<http://www.virec.research.va.gov/RUGs/RUG-MedSAS-OP-FY09-ER.pdf>

Why do we need a CDW Outpatient Workload Query?

- The CDW Outpatient Workload Query is necessary for replication of the mechanism within VistA that filters outpatient workload data to NPCD.
- CDW outpatient data tables are derived directly from local VistA files. Outpatient workload data is not filtered.
 - Cancelled and No-Show appointments in the CDW outpatient tables do not count as workload.
- The CDW Outpatient Workload query excludes these records.

Why do we need a CDW Outpatient Workload Query?

Data Progression from VHA Medical Centers to the CDW and to the Outpatient Tables



Why do we need a CDW Outpatient Workload Query?

Comparison of data flow from VistA between NPCD and CDW

VistA to NPCD	VistA to CDW
HL7 message transmission	Raw extract from VistA
Only contains workload	Most of VistA data is present (i.e. outpatient workload, ancillary, progress notes, pharmacy)
VistA Business Criteria and NPCD Edits applied	VistA Business Criteria and NPCD Edits <u>NOT</u> applied

Why do we need a CDW Outpatient Workload Query?

- To obtain workload data from VistA without NPCD and Outpatient MedSAS files, a standard method to filter outpatient workload from CDW data tables will be required.
- The CDW Outpatient Workload Query advantages are:
 - Meets specific data quality standards
 - Can be customized to meet the needs of VA National Program Offices
 - Can be customized to meet the needs of researchers

Why do we need a CDW Outpatient Workload Query?

The CDW Outpatient Workload Query was designed to produce results that replicate the identification of VistA workload transmitted to NPCD.

- A record is required to satisfy logical business criteria requirements in VistA to be queued and transmitted to NPCD as workload. The required criteria are included in the CDW Outpatient Workload Query.
- **Additional business rule edits are applied prior to loading the data into NPCD and are not included in the CDW Outpatient Workload Query.**

Questions Addressed by Session

- Why do we need a CDW Outpatient Workload Query?
- **What does the CDW Outpatient Workload Query do?**
 - Criteria for workload in VistA
 - Translation to a Relational Database
- What did we learn developing & validating our query?
- How can you use the Query?
- Conclusions

What does the CDW Outpatient Workload Query do?

- The CDW Outpatient Workload Query incorporates all business criteria that classify local VistA entries from the Visit file as workload.
- Ten criteria have been identified as determining factors for workload in the VistA system.

Criteria for workload in Vista

1

- *The record must exist in the Visit file and the patient is not flagged as a test patient.*

2

- *The field 'Hospital Location' in the Visit file cannot be null.*

3

- *In Hospital Location file, the field 'Type' must be 'Clinic.'*

Criteria for workload in VistA

4

- *In Hospital Location file, the field 'Non-Count' cannot be 'Yes.'*

5

- *The field 'Service Category' in the Visit file cannot be 'E' Event (Historical).*

6

- *If the Visit file entry was generated from a scheduled appointment, the appointment status cannot be 'Cancelled by Patient', 'Cancelled by Clinic', or 'No-Show.'*

Criteria for workload in VistA

7

- *There must be at least one corresponding entry in the V Provider file. One provider will need to be designated as primary.*

8

- *There must be at least one corresponding entry in the V CPT file.*

9

- *There must be at least one corresponding entry in the V POV file. One diagnosis will need to be designated as primary. The patient disability classification questions must also be answered when applicable.*

Criteria for workload in VistA

10

- *EXCEPTION: Certain stop codes have been classified as ancillary or occasion of service and do not require a V POV Dx code.*

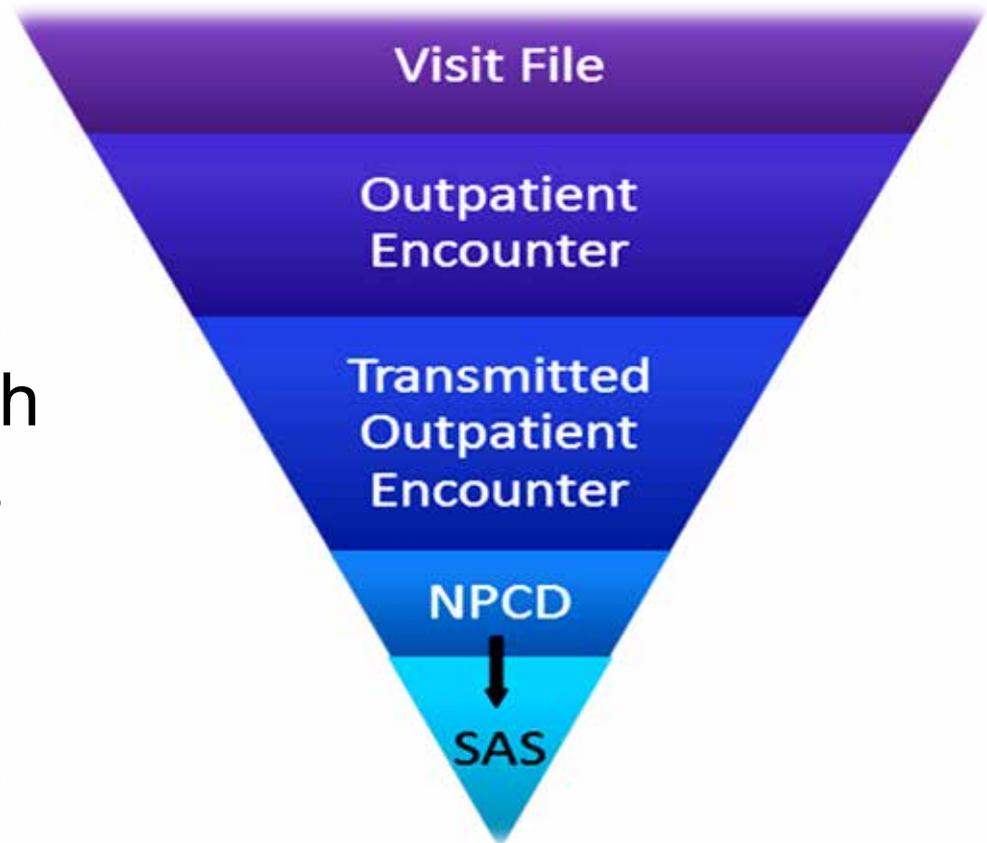
- The stop code exceptions are: 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 126, 127, 128, 144, 145, 146, 149, 150, 151, 152, 153, 160, 170, 317, 421, 703, 999
- These stop codes also do not require disability questions to be answered.

Poll Question

- Have you ever used CDW?
 - Yes! All the time.
 - I am familiar with CDW, but rarely use it.
 - I have never used the CDW.

Translation to a Relational Database

First we set out to ascertain the logic that determines which workload encounters are transmitted to NPCD



Translation to a Relational Database

- In a VistA account, a Fileman query from the Visit file was developed to produce the same records that are transmitted to NPCD from VistA file Transmitted Outpatient Encounter.
- We used the Fileman query results to help develop the CDW Outpatient Workload Query.

Translation to a Relational Database

How do you translate the ten business criteria established within VistA to a relational database?

- Start with the table which contains the patient information needed.
- Find tables which contain *related* information.
- Determine how to link the *related* tables (Table JOIN).
- Determine the conditional statements that define the relationships (WHERE clause).

Translation to a Relational Database

CDW Fact Tables Used in CDW Outpatient Workload Query

Schema	Table Name
Outpat	Visit
Outpat	Vdiagnosis
Outpat	Vprocedure
Outpat	Vprovider
Patient	Patient
Staff	Staff

Translation to a Relational Database

CDW Dimension Tables Used in CDW Outpatient Workload Query

Schema	Table Name
Dim	AppointmentStatus
Dim	CPT
Dim	ICD
Dim	Location

Translation to a Relational Database

Left joins of both fact and dimension tables are used with the `Output.Visit` table to ensure all entries are retrieved.

SQL Code

```
FROM [CDWork].[Output].[Visit]
left join Dim.Location
on Visit.LocationSID=Location.LocationSID
left join Output.VDiagnosis
on Visit.VisitSID=VDiagnosis.VisitSID
join Dim.ICD
on VDiagnosis.ICDSID=ICD.ICDSID
left join Output.VProcedure
on Visit.VisitSID=VProcedure.VisitSID
join Dim.CPT
on VProcedure.CPTSID=CPT.CPTSID
left join Output.VProvider
on visit.VisitSID=VProvider.VisitSID
join SStaff.SStaff
on VProvider.ProviderSID=SStaff.StaffSID
left join DIM.AppointmentStatus
on Visit.AppointmentStatusID=
AppointmentStatus.AppointmentStatusID
```

Translation to a Relational Database

The **WHERE** clause applies the business criteria:

- Location type *must* be clinic
SQL code `Location.LocationType = 'CLINIC'`
- Clinic cannot be flagged as NON-COUNT
SQL code `Location.NonCountClinicFlag = 'N'`
- Service Category of **E**vent cannot be HISTORICAL
SQL code `Visit.ServiceCategory <> 'E'`

Translation to a Relational Database

The **WHERE** clause applies the business criteria:

- Appointment Status *must* be Inpatient Appointment or Checked Out

SQL code (AppointmentStatus.AppointmentStatus = 'INPATIENT APPOINTMENT'
OR AppointmentStatus.AppointmentStatus = 'CHECKED OUT')

- A CPT code is required

SQL code VProcedure.CPTSID is not NULL

- A **P** primary provider is required

SQL code VProvider.PrimarySecondary = 'P'

Translation to a Relational Database

The **WHERE** clause applies the business criteria:

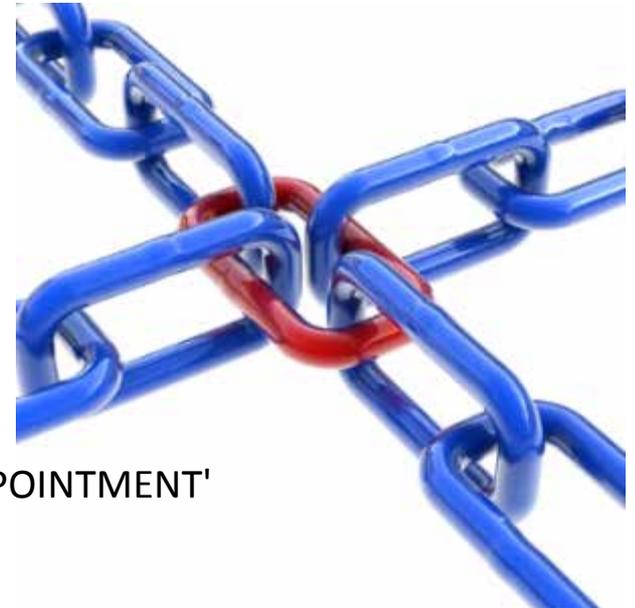
- A **P** primary diagnosis is required on all entries unless the location has a specific stop code.

```
SQL code      (VDiagnosis.PrimarySecondary = 'P'  
              or Visit.PrimaryStopCode in ('104', '105', '106', '107', '108',  
              '109', '110', '111', '112', '113', '114', '115', '126', '127',  
              '128', '144', '145', '146', '149', '150', '151', '152', '153',  
              '160', '170', '317', '421', '703', '999'))
```

Translation to a Relational Database

SQL Code WHERE Clause

```
WHERE VISIT.Sta3n = '531'  
and Visit.VisitDateTime BETWEEN  
    '2012-01-01 00:00:00' and '2012-03-31 23:59'  
and Location.LocationType = 'CLINIC'  
and Location.NonCountClinicFlag = 'N'  
and Visit.ServiceCategory <> 'E'  
and VProcedure.CPTSID is not NULL  
and (AppointmentStatus.AppointmentStatus = 'INPATIENT APPOINTMENT'  
OR AppointmentStatus.AppointmentStatus = 'CHECKED OUT')  
and VProvider.PrimarySecondary = 'P'  
and(VDiagnosis.PrimarySecondary = 'P'  
or Visit.PrimaryStopCode in ('104', '105', '106', '107', '108',  
    '109', '110', '111', '112', '113', '114', '115', '126', '127',  
    '128', '144', '145', '146', '149', '150', '151', '152', '153',  
    '160', '170', '317', '421', '703', '999'))
```



Questions Addressed by Session

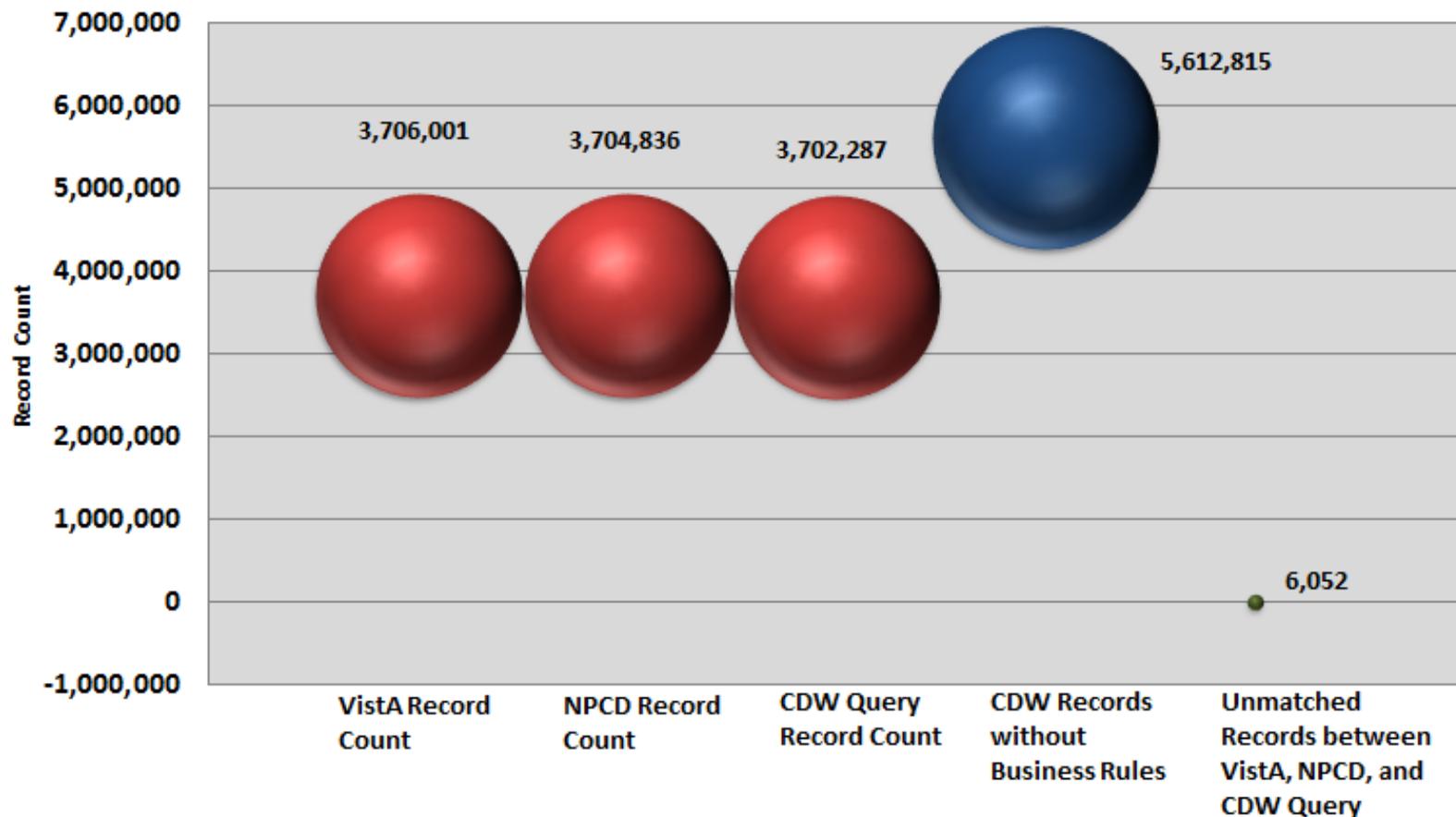
- Why do we need a CDW Outpatient Workload Query?
- What does the CDW Outpatient Workload Query do?
 - Criteria for workload in VistA
 - Translation to a Relational Database
- What did we learn developing & validating our query?
- How can you use the Query?
- Conclusions

What did we learn developing & validating our query?

- A comparison between VistA, NPCD, and CDW was conducted for eleven VistA systems.
 - Date range used 1/1/2012-3/31/2012
- A unique identifier, called Visit ID in the VistA Visit File, is present in all three sources for comparative purposes.

What did we learn developing & validating our query?

Comparison of Records Between VistA, NPCD, and CDW



What did we learn developing & validating our query?

Changes to VistA Files that Affect CDW Queries

- CDW queries may not reflect VistA field values at the time of record entry to VistA
- Data queried from CDW contains current values of core files (i.e., Hospital Location).
 - Local VistA changes in Non-Count field from 'No' to 'Yes'.

How did we validate the Query and what did we learn?

There are NPCD data load edits that need to be applied to workload and addressed in the CDW environment.

- Invalid means test
- Record contains expired CPT code
- Employee records missing registration fields
- Records entered after date of patient's death
- Eligibility and POW status not updated

How did we validate the Query and what did we learn?

There are some fields in VistA that are missing in CDW.

- CPT modifiers
- Patient disability classification questions

Questions Addressed by Session

- Why do we need a CDW Outpatient Workload Query?
- What does the CDW Outpatient Workload Query do?
 - Criteria for workload in VistA
 - Translation to a Relational Database
- What did we learn developing & validating our query?
- **How can you use the Query?**
- Conclusions

How can you use the Query?

- The query provides a **foundation** to identify workload records.
- Using SQL will enable users to customize the data for their specific needs.
- By using VistA data and replicating the business logic, the CDW Outpatient Workload Query provides more accurate data than record management performed through HL7 messaging.

Questions Addressed by Session

- Why do we need a CDW Outpatient Workload Query?
- What does the CDW Outpatient Workload Query do?
 - Criteria for workload in VistA
 - Translation to a Relational Database
- What did we learn developing & validating our query?
- How can you use the Query?
- **Conclusions**

Conclusions

- The CDW Outpatient Workload Query validation was successful. The query matches +99% of the baseline data from VistA.
- The unmatched entries were resolved and found to be a result of how the data is recorded at each source.
- The CDW Outpatient Workload Query will improve data quality by eliminating errors inherent in the transmission process.

Conclusions

- NPCD edits will need to be included in future versions of the CDW Outpatient domain data
- The CDW Outpatient Workload Query will need to be expanded to include ALL elements of the Outpatient medical SAS datasets
- Allows VHA to improve data quality beyond the edits implemented for NPCD and the Outpatient Medical SAS datasets

Questions and Comments

Trinity Hall Trinity.Hall@va.gov

Laureen Kennedy Laureen.Kennedy@va.gov

Data Quality Analysis mail group:
[VHA 10P2 OIA DQA](#)

Data Quality Website:
http://vaww.vhadataquality.va.gov/index.php?option=com_content&view=article&id=117&Itemid=452&lang=en

