Database & Methods Cyberseminar Series

Overview of VA Data, Information Systems, National Databases & Research Uses

October 3, 2016
Maria Souden, MSI, PhD
Associate Director, VA Information Resource Center
What VIReC does

- Develop knowledge about VA data
- Share knowledge with VA researchers
  - Intranet website
  - VHA Data Portal (collaboration with NDS, VINCI, & Data Quality)
  - Cyberseminars (hosted by CIDER)
  - Guides, documentation, and tutorials
  - HelpDesk (virec@va.gov)
Informational seminars to help VA researchers access and use VA databases.

Topics

• VA data sources & data access systems
• Application of VA data to research and quality improvement questions
• Limitations of secondary data use
• Resources to support VA data use
## FY ‘17 Database & Methods Schedule

First Monday of the month*  |  1:00pm-2:00pm ET

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/3/16</td>
<td>Overview of VA Data &amp; Research Uses</td>
</tr>
<tr>
<td>11/7/2016</td>
<td>Requesting Access to VA Data</td>
</tr>
<tr>
<td>12/5/2016</td>
<td>VA Healthcare Utilization</td>
</tr>
<tr>
<td>1/9/2017*</td>
<td>VA Medicare Data (VA/CMS)</td>
</tr>
<tr>
<td>2/6/2017</td>
<td>Measuring &amp; Assessing Outpatient Care</td>
</tr>
<tr>
<td>3/6/2017</td>
<td>Mortality Ascertainment &amp; Cause of Death</td>
</tr>
<tr>
<td>4/3/2017</td>
<td>Assessing Race &amp; Ethnicity</td>
</tr>
<tr>
<td>5/1/2017</td>
<td>Co-morbidity Measures Using VA &amp; CMS Data</td>
</tr>
<tr>
<td>5/22/2017*</td>
<td>Advanced Topics in Co-morbidity</td>
</tr>
<tr>
<td>6/5/2017</td>
<td>Pharmacy Data</td>
</tr>
<tr>
<td>7/10/2017*</td>
<td>CAPRI/VistAWeb for EHR Access</td>
</tr>
<tr>
<td>9/11/2017*</td>
<td>CDW microbiology, lab, &amp; pharmacy domains</td>
</tr>
</tbody>
</table>

*Schedule shifts by one week in event of VA holiday.

Visit our Education page for more information & registration links. [www.virec.research.va.gov](http://www.virec.research.va.gov)
Session 1: Overview of VA Data, Information Systems, National Databases & Research Uses
Poll #1: Your role as a data user

• What is your role in research and/or quality improvement?
  • Research investigator
  • Methodologist
  • Data manager, analyst, or programmer
  • Project coordinator
  • Other – please describe via the Q&A function
Poll #2: Your experience with VA data

How many years of experience do you have working with VA data?

- One year or less
- More than 1, less than 3 years
- At least 3, less than 7 years
- At least 7, less than 10 years
- 10 years or more
The purpose of this cyberseminar is to provide an introduction to VA data and resources for using it effectively.

Today’s objective
Session roadmap

- Introduction to VA data
- Highlights of commonly used VA data sources
- Platforms & tools for using data
- Where to go for more help
Session roadmap

• Introduction to VA data
• Highlights of commonly used VA data sources
• Platforms & tools for using data
• Where to go for more help
Activities that generate VA data

• Administrative / Operations activities
• Clinical care / Patient records
• Financial recordkeeping
• Veteran population surveys
• Patient care outside the VA system
Veterans Health Information Systems & Technology Architecture (VistA)

- Integrated system of software applications.
- Directly supports patient care at Veterans Health Administration (VHA) healthcare facilities.
- Source of data for many VHA datasets or databases of interest to researchers.
Commonly used VA data sources

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA/CMS Data for Research
Commonly used VA data sources

- **EHR/VistA**
  - VA Corporate Data Warehouse (CDW)

- **VA Medical SAS (MedSAS)**

- **VA Vital Status File**

- **VA MCA National Clinical Data**

- **VA Pharmacy Data**

- **VA CMS Data for Research**

Date: 10/2016
Commonly used VA data sources

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA CMS Data for Research

Utilization and Status
Commonly used VA data sources

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA CMS Data for Research

Clinical Care Activities & Costs
Commonly used VA data sources

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA Vital Status File
- VA CMS Data for Research

Non-VA Data

10/2016
Session roadmap

- Introduction to VA data
- Highlights of commonly used VA data sources
- Platforms & tools for using data
- Where to go for more help
- Questions
Highlight: Corporate Data Warehouse

Data steward: National Data Systems (NDS)

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA CMS Data for Research

10/2016
CDW origin & content

>130 VistA Systems

Region 1 Data Warehouse

Region 2 Data Warehouse

Region 3 Data Warehouse

Region 4 Data Warehouse

CDW (60+ domains)

Patient
Staff
Allergy
Inpatient
Outpatient
Inpatient Rx
Outpatient Rx
Lab Chemistry
Lab Microbiology
Vital Signs
etc.

Minute by minute transactions added – no changes

Nightly updates to architected database - no change to content.
CDW… things to know

- Data straight from VistA
  - Not curated or transformed
- CDW is a relational database, *not* a dataset
- For SAS users, requires some conceptual shifts and new skills
- Relatively new and still evolving
- CDW SharePoint Site:
  [https://vaww.cdw.va.gov/Pages/CDWHome.aspx](https://vaww.cdw.va.gov/Pages/CDWHome.aspx) (VA Intranet)
CDW roles

**BISL**
- Data Architects
- Business Intelligence Service Line
- Builds the data warehouse that organizes these data

**NDS**
- Data Steward
- National Data Systems
- Grants access to CDW data

**VINCI**
- Data Provisioning
- VA Informatics & Computing Infrastructure
- Creates subsets of CDW data for users who have been granted access

**VIReC**
- Data Knowledge
- VA Information Resource Center
- Provides documentation & education to help understand and use the data

10/2016
VIReC resources for CDW data

http://vaww.virec.research.va.gov/CDW/Documentation.htm
(VA Intranet)
Part 3. Clinical Context

According the Outpatient Domain Release Document [28], the primary fact table \textit{Outpat.Visit} contains records from three VistA files: Visit (S#0000010) file, Outpatient Encounter (S#09.68) file and the Patient (2) file. The VistA file Visit (S#0000010) is a collection of care and non-care records related to patients and non-patients that

Part 4. Table and Column Descriptions

Part 4 describes each table in the Outpatient 2.1 Domain and selected columns within each table. It is organized by table. The name of the table is presented in large font with a description of its general content immediately following. After each table description, selected columns within that table are described. Table 1 summarizes the various parts of the column descriptions.

Table 1. Content of Column Descriptions

Part 5. Primary and Foreign Key Connections Query

Once you have been granted access to CDW data, this query can be used to generate an up-to-date list of primary and foreign keys associated with this domain.

Part 6. Example SQL Code for Basic Research Questions

This section contains a series of simple research questions that are answered by applying basic SQL queries to the data in this domain. The three research questions include:

- How does \textit{Outpat.Visit} compare to \textit{Outpat.Workload}, in terms of number of records?
- What appointment status is associated with visits at my station last month?
- What are the 10 most common primary diagnoses for outpatient encounters at my station last month?

Question 1. How does \textit{Outpat.Visit} compare to \textit{Outpat.Workload}, in terms of number of records?

\begin{verbatim}
SELECT COUNT (VisitSID) as TotalNumVisitsInJan
FROM CDW.Visit
WHERE VisitDateTime <= CONVERT (DATETIME, '1/1/2015')
AND VisitDateTime <= CONVERT (DATETIME, '1/31/2015');
\end{verbatim}

TotalNumVisitsInJan

\begin{verbatim}
17,652,832
\end{verbatim}

\begin{verbatim}
SELECT COUNT (VisitSID) as TotalWorkloadVisits
FROM CDW.Workload;
\end{verbatim}
Highlight: VA MedSAS Data
Data steward: National Data Systems (NDS)
MedSAS… things to know

• Data have been summarized and had business rules applied
  • Inpatient discharges at the stay level (episode of care)
  • Outpatient information at the visit and encounter level

• Include diagnostic, treatment, and some demographic information from the electronic health record
  • Demographic information current at time of encounter or stay

• Includes data about care in non-VA facilities paid for by VA
MedSAS inpatient overview

Patient Treatment File (PTF)
Austin Information Technology Center (AITC)
National Data Systems (NDS)
MedSAS inpatient data elements

Examples

- Patient demographics
- ICD-9 diagnosis codes
- ICD-9 procedure/surgery codes
- Discharge dates
- Treating specialty
MedSAS outpatient overview

National Patient Care Database (NPCD)
National Data Systems (NDS)
MedSAS outpatient data elements

Examples

- Patient demographics
- ICD-9 diagnosis codes
- CPT-4 procedure codes
- Event dates
- Provider specialty
Coming soon! MedSAS Replacement data view

- NPCD Database to be discontinued.

- MedSAS outpatient files will become available as SQL views in CDW
  - Current variables recreated exactly
  - Will look the same to the end user, just in SQL

- Being tested to make sure workload logic applied correctly

- Availability and access information to come
VIReC resources for MedSAS data

(VA Intranet)
2.1 Background

VA health care facility staff members across the United States and Puerto Rico enter patient care information into Veterans Health Information Systems and Technology Architecture (VistA), the main information collection and data management system in the Veterans Health Administration (VHA). A select set of data collected during patient health care occasions are entered in local health care facilities VistA systems and then electronically transferred from each local VistA system upon discharge of a patient or upon collection of quarterly census data only to a centralized Patient Treatment Module (PTM) environment of the Veterans Information Technology and Architecture (VITA) program or the Active Information Technology (AIT) program.

3.1 List of Variables & their Dataset Locations

This section lists variables and their dataset locations for the fiscal year (FY) 2014 VHA Medical SAS Inpatient Acute Care Main, Bed Section, Procedure and Surgery Datasets. An X in the frequency column indicates VIReC provides additional summary information for this variable in VIReC Select Variable Frequencies: Medical SAS Datasets, FY2000–2014 [16].

Table 5 provides a short definition of each variable and reports their presence in the FY2014 Inpatient Acute Care Datasets.

Table 5. FY2014 VHA Medical SAS Inpatient Variables and their Dataset Locations

<table>
<thead>
<tr>
<th>SAS Name</th>
<th>Definition</th>
<th>Main</th>
<th>Bed Section</th>
<th>Surgery</th>
<th>Procedure</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABO</td>
<td>Number of days a patient was out of the hospital on pass during an inpatient stay</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADMITDAY</td>
<td>Date of admission of the inpatient stay</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Variable Name: ABO

Definition: Number of days a patient was out of the hospital on pass during an inpatient stay

Remarks: This variable is used in the calculation of the LS (number of bed days for the entire episode of care) variable. Chapter 10, “Absences,” of the VHA Manual of Policies (Ad-1): Operations: Part 1, Medical Administration Activities [17] provides guidelines for granting of authorized absences. Section 10.6.8 of this manual indicates that an authorized absence for hospital patients may not exceed 96 hours. The 96-hour limit applies to the variables PASS and PASSB. This section suggests that ABO is reserved for absences greater than 96 hours.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Numeric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Print Format</td>
<td>None</td>
</tr>
<tr>
<td>Label</td>
<td>ABSENT BED OCCUPANT DAYS</td>
</tr>
<tr>
<td>Variable Availability in Medical SAS Datasets: Fiscal Years (FY)</td>
<td>Acute Main, FY1970 – to date</td>
</tr>
<tr>
<td>VistA Data Source</td>
<td>PTI (45) file, ASIH field</td>
</tr>
</tbody>
</table>
Highlight: VA Vital Status File (VSF)
Data steward: National Data Systems (NDS)
VA Vital Status File… things to know

- Includes death and birth dates from multiple sources of mortality information
  - VA Workload Data (PTF and Fee Basis)
  - Beneficiary Identification and Resource Locator System (BIRLS) Death File (from VBA)
  - Medicare Vital Status File
  - Monthly Social Security Administration Death Master File

- Two files - Master and Mini
# VA Vital Status File data elements

## VSF Master File
- Death dates from:
  - VA & CMS enrollment & utilization
  - VBA & Social Security benefits
- Last healthcare utilization dates from multiple sources

## VSF Mini File
- Applies algorithm to determine “best”
  - Death date
  - Birth date
  - Gender
- Last utilization
- Presumed living indicator
Highlight: MCA National Clinical Data

Data steward: VHA Managerial Cost Accounting Office

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA CMS Data for Research

10/2016
MCA national clinical data… things to know

- Available as 26 National Data Extracts (NDEs), classified as Core, Clinical, Financial, Program Activity
  - See VHA Data Portal for full list
- Four Clinical NDEs most relevant to health services research
  - Lab tests & costs (LAB)
  - Lab test results (LAR)
  - Pharmacy prescriptions (PHA)
  - Radiology procedures (RAD)
VA MCA clinical NDE data elements

Examples

- Date
- Test results
- Ordering provider
- Costs
Resources for MCA NDEs

VIReC MCA NDEs Documentation
(VA Intranet)

MCAO National Data Extracts & Reporting Information
http://vaww.dss.med.va.gov/nationalrptg/nr_extracts.asp
(VA Intranet)

Health Economics Resource Center (HERC)
http://vaww.herc.research.va.gov/include/page.asp?id=managerial-cost-accounting
(VA Intranet)
Highlight: VA pharmacy data
Data stewards: MCAO, PDM, NDS

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data
- VA CMS Data for Research
Sources of data on medications dispensed in VA

MCA Pharmacy NDE
  • Data Steward: MCAO

Pharmacy Benefits Management
  • Data Steward: PBM

Two CDW Domains
  • Data Steward: NDS
  • Bar Code Medication Administration (BCMA) Domain (inpatient pharmacy)
  • Outpatient Pharmacy Domain
VA pharmacy data elements

Examples

• Generic drug name
• Days supply
• VA drug class
• Cost
Resources for pharmacy data

http://vaww.virec.research.va.gov/Pharmacy/Overview.htm

(VA Intranet)
Highlight: VA/CMS Data for Research
Data steward: VIReC

- VA Corporate Data Warehouse (CDW)
- VA Medical SAS (MedSAS)
- VA Vital Status File
- VA MCA National Clinical Data
- VA Pharmacy Data

VA/CMS Data for Research

10/2016
VA/CMS - types of data

- Medicare & Medicaid enrollment and claims data
- Medicare Current Beneficiary Survey (MCBS) data
- Patient assessment data
- Healthcare Effectiveness Data and Information Set (HEDIS) - utilization in Medicare managed care
- United States Renal Data System (USRDS) data
VA/CMS data elements

- Examples
  - Medicare and Medicaid enrollment
  - ICD-9 diagnosis codes
  - ICD-9 and CPT4 procedure codes
  - Charges and payments
  - Veteran flag
Resources for VA/CMS Data

http://vaww.virec.research.va.gov/Index-VACMS.htm
(VA Intranet)
Session roadmap

• Introduction
• Highlights of commonly used VA data sources
• Platforms & tools for using data
• Where to go for more help
VA Informatics and Computing Infrastructure

- Secure data environment
- High-speed processing platform
- Free data management and analytic tools
- VINCI services

VINCI Central Website
http://vaww.vinci.med.va.gov/vincicentral/ (VA Intranet)
AICT (Austin Information Technology Center)

- Centralized computer processing center (mainframe)
- Available data
  - MedSAS datasets
  - VA Vital Status file
- Learn More…
  - VHA Data Portal AICT Mainframe page
    [http://vaww.vhadataportal.med.va.gov/DataSources/AITCMainframe.aspx](http://vaww.vhadataportal.med.va.gov/DataSources/AITCMainframe.aspx) (VHA Intranet)
  - AICT Mainframe Website
    [https://vaww.sde.portal.va.gov/sites/oe/Pages/Default.aspx](https://vaww.sde.portal.va.gov/sites/oe/Pages/Default.aspx) (VHA Intranet)
# CAPRI & VistAWeb…things to know

*Read-only access to patient electronic health records (EHRs)*

<table>
<thead>
<tr>
<th>CAPRI</th>
<th>VistAWeb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses desktop computer software application</td>
<td>VA Intranet web portal</td>
</tr>
<tr>
<td>Can see patient record one site at a time</td>
<td>Can see patient record for all sites at once</td>
</tr>
<tr>
<td>Includes search function</td>
<td>VistA Imaging available</td>
</tr>
<tr>
<td>Direct access to VistAWeb</td>
<td></td>
</tr>
</tbody>
</table>

10/2016
Session roadmap

- Introduction
- Highlights of commonly used VA data sources
- Platforms & tools for using data
- Where to go for more help
Welcome to the VHA Data Portal

The VHA Data Portal provides a one-stop-shop for data users’ needs. The portal integrates information from multiple sources into a single location, providing a comprehensive knowledge base and the functionality needed for efficient and effective research.

Upcoming Events

VA ReCyberseminar

Next event: October 2016

VA network only

Combined resource (with NDS and VINCI)

Data source & access information

http://vaww.vhadataportal.med.va.gov

www.virec.research.va.gov

Public-facing VIReC overview
Publications & events
Cyberseminars

vaww.virec.research.va.gov

VA network only
Data news
Resources for working with data

10/2016
VIReC options for specific questions

**HSRData Listserv**
- Community knowledge sharing
- ~1,200 VA data users
- Researchers, operations, data stewards, managers

**HelpDesk**
- Individualized support
  - [virec@va.gov](mailto:virec@va.gov)
  - (708) 202-2413

10/2016
## Quick links for VA data resources

### Quick Guide: Resources for Using VA Data

### VIReC: [http://vaww.virec.research.va.gov/Index.htm](http://vaww.virec.research.va.gov/Index.htm) (VA Intranet)


### CDW: [https://vaww.cdw.va.gov/Pages/CDWHome.aspx](https://vaww.cdw.va.gov/Pages/CDWHome.aspx) (VA Intranet)

### Archived cyberseminar: What can the HSR&D Resource Centers do for you?
Take-aways

- VA has many data sources that can be used to address research & QI questions
- Data are complex and distributed
- Use resources VIReC and others offer to
  - help you navigate data sources
  - connect you with peers, experts & additional info
Contact information

Maria Souden, MSI, PhD
Associate Director
VA Information Resource Center
Hines VA Hospital
virec@va.gov
708-202-2413
Next session:
November 7, 2016

Database & Methods Cyberseminar Series

Requesting Access to VA Data for Research and Quality Improvement

Linda Kok, MA
VA Information Resource Center