

VIReC Database & Methods Cyberseminar Series

Clinical Epidemiology Research using National MCA and CDW Laboratory Data: Perspectives from the Frontline

May 2, 2016

Presented by:

Benjamin Bowe, MPH

Yan Xie, MPH

Ziyad Al-Aly, MD

Acknowledgements

- Anne O'Hare, MD
 - VA HSR&D Center of Excellence & University of Washington
- Adam Batten
 - VA Puget Sound Health Care System
- Jeff Todd-Stenberg
 - VA Puget Sound Health Care System
- Daniel Bertenthal
 - San Francisco VA Medical Center

Session Outline

- Overview of laboratory data in the VA Managerial Cost Accounting (MCA) National Data Extracts (NDEs)
- Our experience using VA MCA NDE laboratory data
- Overview of laboratory data in the VA Corporate Data Warehouse (CDW)
- Our experience using VA CDW laboratory data

Session Outline

- Overview of laboratory data in the VA Managerial Cost Accounting (MCA) National Data Extracts (NDEs)
- Our experience using VA MCA NDE laboratory data
- Overview of laboratory data in the VA Corporate Data Warehouse (CDW)
- Our experience using VA CDW laboratory data

Poll Question #1: I am interested in VA data primarily due to my role as _____.

- Research investigator
- Data manager
- Project coordinator
- Program specialist or analyst
- Other (specify)

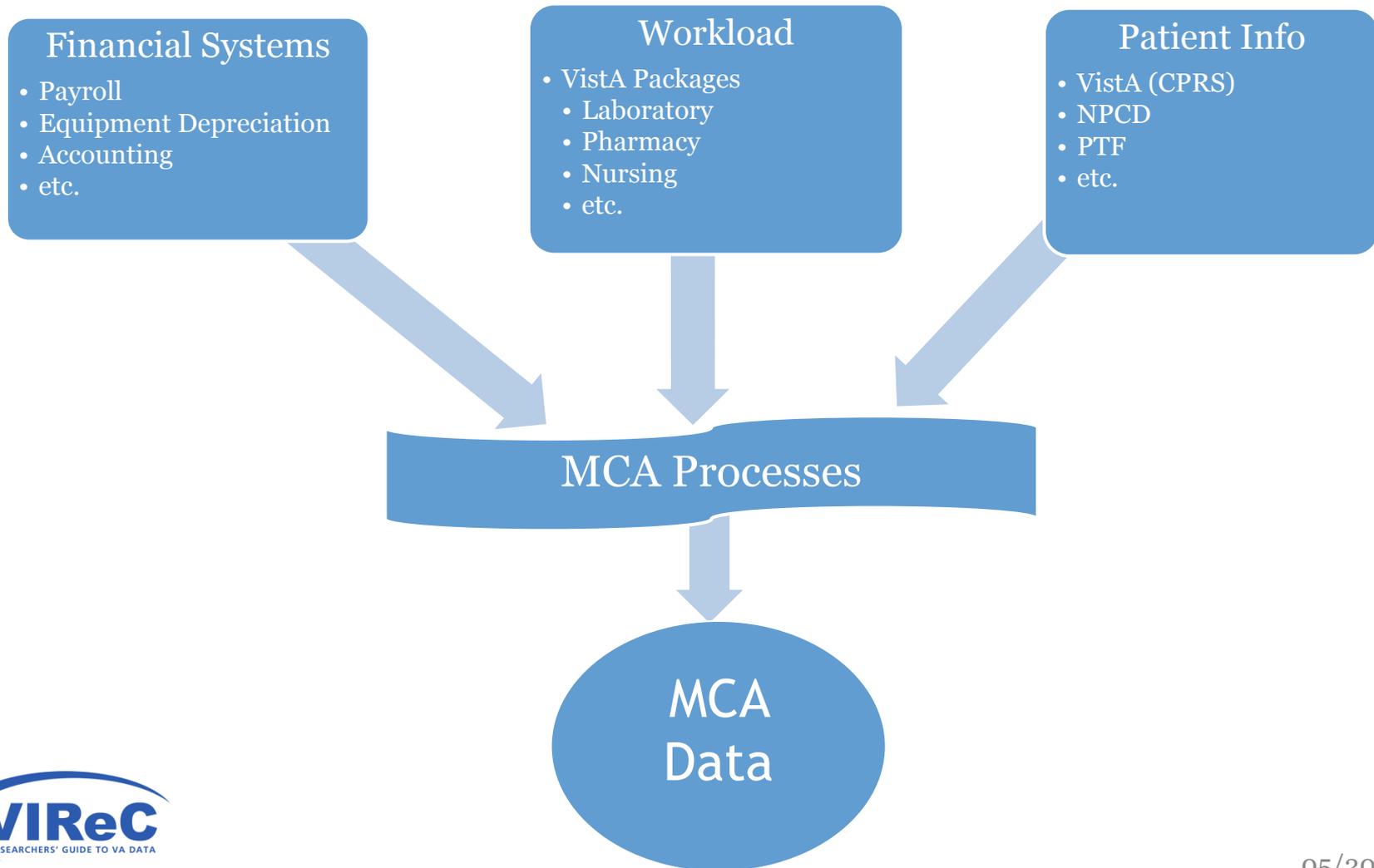
Poll Question #2: Have you worked with lab data in MCA NDEs?

- Yes
- No

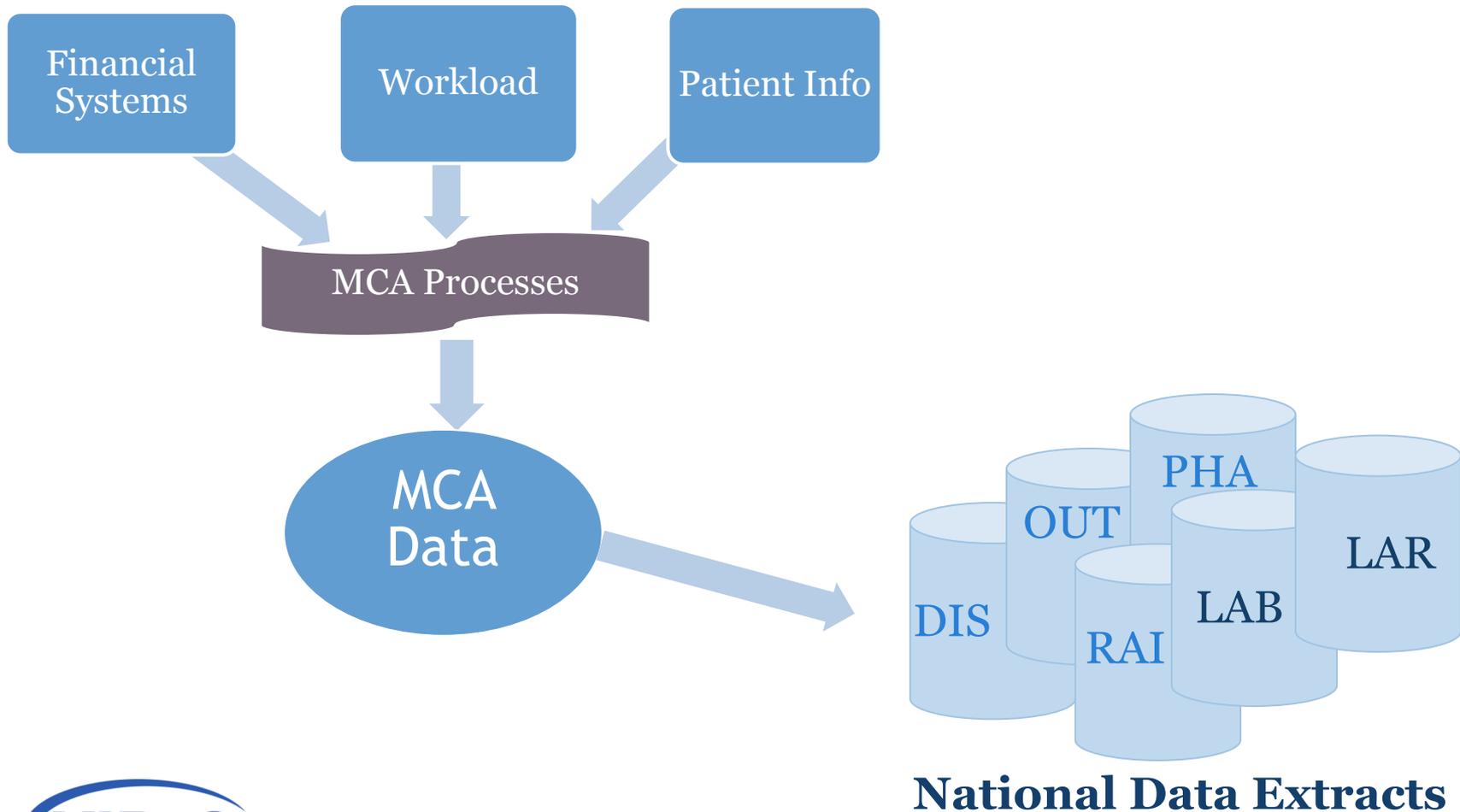
MCA Overview

- What is MCA?
 - VA's Managerial Cost Accounting System
 - Formerly Decision Support System (DSS)
- What is its primary purpose?
 - MCA information supports process and performance improvement by measuring quality of care, clinical outcomes, and financial impact

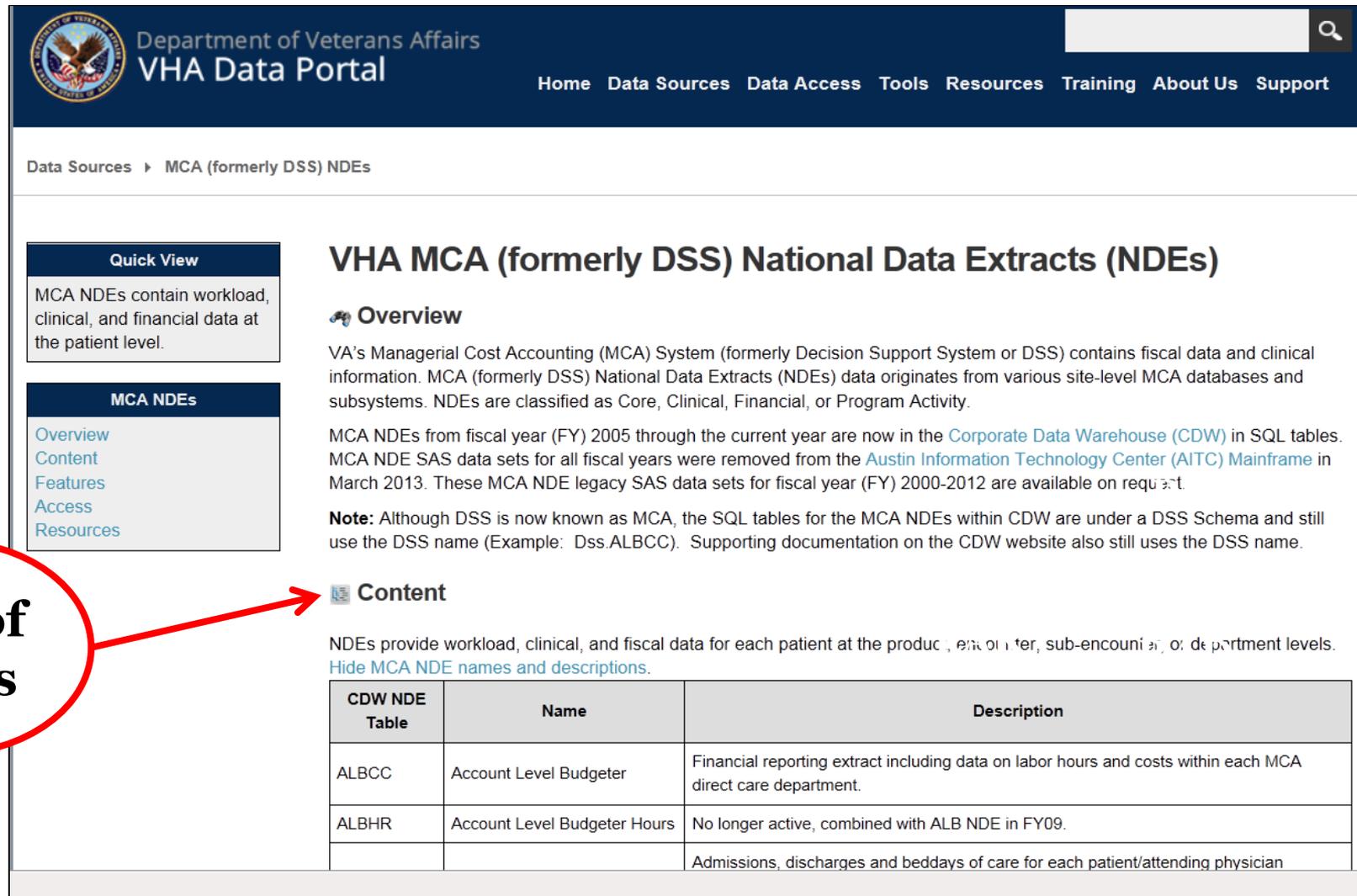
MCA Source Data



MCA National Data Extracts (NDEs)



VA Managerial Cost Accounting System National Data Extracts (MCA NDEs)



Department of Veterans Affairs
VHA Data Portal

Home Data Sources Data Access Tools Resources Training About Us Support

Data Sources > MCA (formerly DSS) NDEs

Quick View

MCA NDEs contain workload, clinical, and financial data at the patient level.

MCA NDEs

[Overview](#)
[Content](#)
[Features](#)
[Access](#)
[Resources](#)

VHA MCA (formerly DSS) National Data Extracts (NDEs)

Overview

VA's Managerial Cost Accounting (MCA) System (formerly Decision Support System or DSS) contains fiscal data and clinical information. MCA (formerly DSS) National Data Extracts (NDEs) data originates from various site-level MCA databases and subsystems. NDEs are classified as Core, Clinical, Financial, or Program Activity.

MCA NDEs from fiscal year (FY) 2005 through the current year are now in the [Corporate Data Warehouse \(CDW\)](#) in SQL tables. MCA NDE SAS data sets for all fiscal years were removed from the [Austin Information Technology Center \(AITC\) Mainframe](#) in March 2013. These MCA NDE legacy SAS data sets for fiscal year (FY) 2000-2012 are available on request.

Note: Although DSS is now known as MCA, the SQL tables for the MCA NDEs within CDW are under a DSS Schema and still use the DSS name (Example: Dss.ALBCC). Supporting documentation on the CDW website also still uses the DSS name.

Content

NDEs provide workload, clinical, and fiscal data for each patient at the product, service center, sub-encounter, or department levels. [Hide MCA NDE names and descriptions.](#)

CDW NDE Table	Name	Description
ALBCC	Account Level Budgeter	Financial reporting extract including data on labor hours and costs within each MCA direct care department.
ALBHR	Account Level Budgeter Hours	No longer active, combined with ALB NDE in FY09.
		Admissions, discharges and beddays of care for each patient/attending physician

**List of
NDEs**

Laboratory National Data Extracts (NDEs)

Clinical NDEs

- LAB
 - Workload and costs
 - Test-level records
- LAR
 - Laboratory results for a defined list of tests (currently 95)
 - Test-level records

MCA NDEs

- Update schedule: Monthly or quarterly
- Cumulative year-to-date
- Lab data from FY2000 (LAR) or 2002 (LAB)

MCA NDE Data Formats

- SQL tables
 - SQL file for all LAR results, across fiscal years and VISNs
 - Available for fiscal years 2000-2004 and fiscal year 2005 onward
 - Resides in Corporate Data Warehouse (CDW)

The Laboratory Result Data

- Contains information on 95 different labs
 - http://vaww.dss.med.va.gov/programdocs/pd_clinictop.asp (VA Intranet only)
- Labs were sequentially added as time went on, and therefore data availability varies by lab
 - Data starts at year of addition

Strengths and limitations of MCA Laboratory data

- **Strengths**
 - Easy for end-user (single identifier DSSLARNO for each test)
- **Limitations**
 - Incomplete capture of all relevant lab tests at all medical centers
 - Algorithm for mapping process not always readily available, especially for earlier years
 - Possible for contamination of results (ex: urine creatinine in the serum creatinine data)

LOINC Codes

- LOINC: Logical Observation Identifier Names and Codes
 - <http://www.loinc.org>
- Standardized, highly specific - Identifies test, method of analysis, specimen source
 - Multiple Codes for what you may think is one lab test
- Lab results (LAR) for DSS records pulled based on LOINC, implemented nationwide back to FY 2009 onward, available in LAB from FY 2013 onward

Cautions

- DSSLARNO may not include all the LOINC you want
- For higher DSSLARNO the data often does not go back all the way to FY2000
- May have patient identifiers in MCA that are not in other datasets

Session Outline

- Overview of laboratory data in the VA Managerial Cost Accounting (MCA) National Data Extracts (NDEs)
- **Our experience using VA MCA NDE laboratory data**
- Overview of laboratory data in the VA Corporate Data Warehouse (CDW)
- Our experience using VA CDW laboratory data

Examples of LAR TESTS used in our work

- Serum creatinine—DSSLARNO 31
- Microalbumin-to-creatinine ratio—DSSLARNO 56
- High Density Lipoprotein Cholesterol (HDL-C)-
DSSLARNO 28

Serum Creatinine

- A measure of kidney function
- Transformed into estimated Glomerular Filtration Rate (eGFR) using an equation (factors in race, age, gender)
 - Often stratify models by this, or use it for cohort inclusion/exclusion criteria
 - Used for various CKD progression outcomes

Considerations for Data Cleaning of Serum Creatinine

- Consider only clinically viable values
 - Cleaned as valid values are >0.3 and <20 MG/DL
- Text results are considered missing
- Outpatient and/or inpatient depending on what we are doing
 - Inpatient values aren't very stable, can highly influence assessment of longitudinal trend

HDL-C And Risk Of CKD Progression

- Relationship between high density lipoprotein cholesterol (HDL-C) and risk of chronic kidney disease (CKD) progression (as change in serum creatinine and eGFR)
- HDL-C, LDL-C, triglycerides, and serum creatinine from MCA lab data

Bowe B, Xie Y, Xian H, et al. Low levels of high-density lipoprotein cholesterol increase the risk of incident kidney disease and its progression. *Kidney Int.* 2016;89: 886–896.

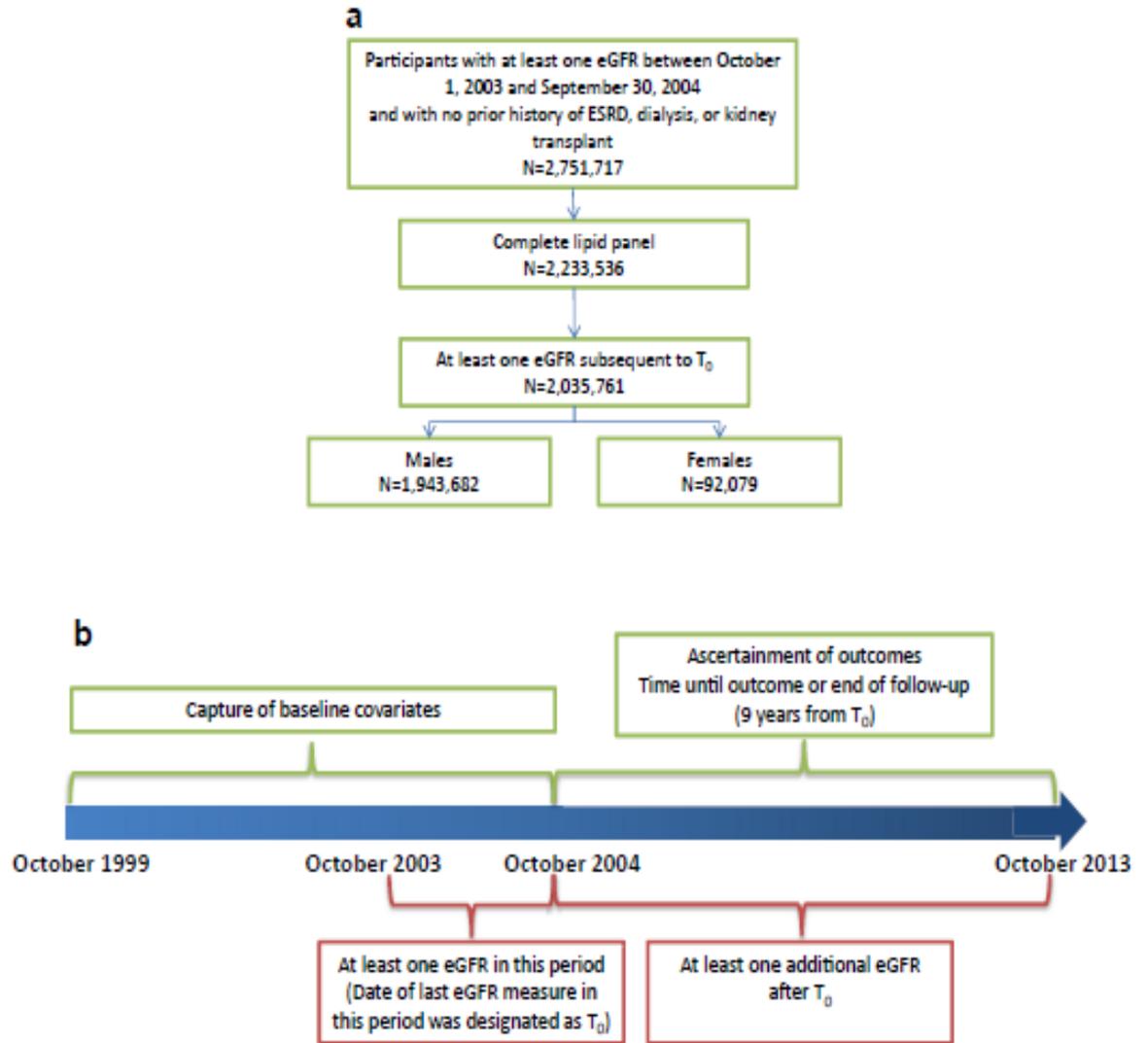


Figure 3 | (a) Flow diagram of cohort assembly. (b) Timeline of cohort selection.

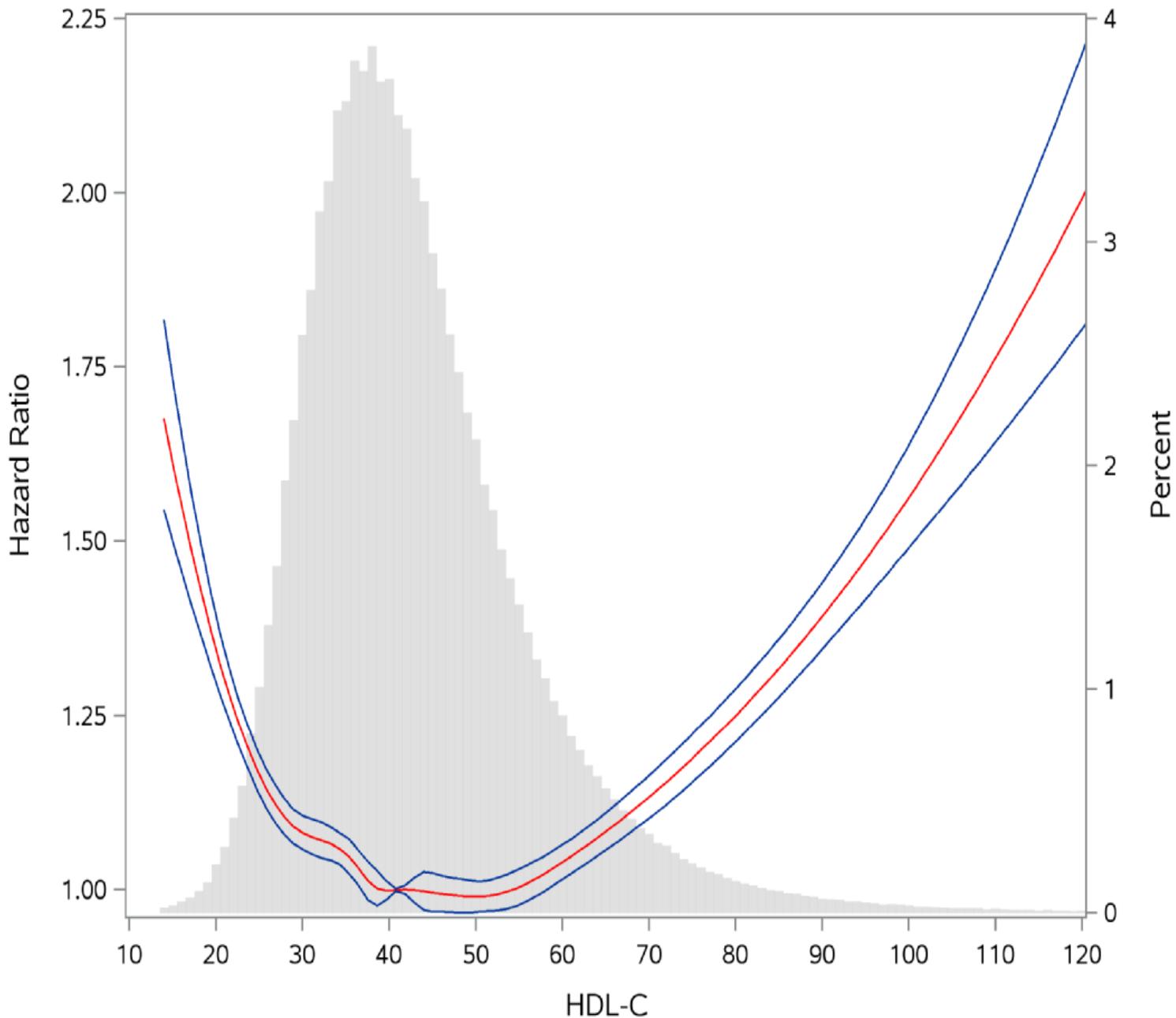


Figure 2 | (a) Cubic spline analyses of risk of doubling of serum creatinine by high-density lipoprotein cholesterol (HDL-C) level (median HDL-C as reference) with HDL-C probability distribution histogram represented by gray bars.

eGFR Trajectories Of Those Entering CKD Stage 4

- CKD stage 4 is the stage before a person enters End Stage Renal Disease (ESRD), which requires dialysis/kidney transplant
- Study aim was to investigate the eGFR trajectory into CKD stage 4, the factors associated with each trajectory, and how outcomes differed by trajectory

Xie, Y., Bowe, B., Xian, H., Balasubramanian, S., & Al-Aly, Z. (2016). Estimated GFR Trajectories of People Entering CKD Stage 4 and Subsequent Kidney Disease Outcomes and Mortality. American Journal of Kidney Diseases.

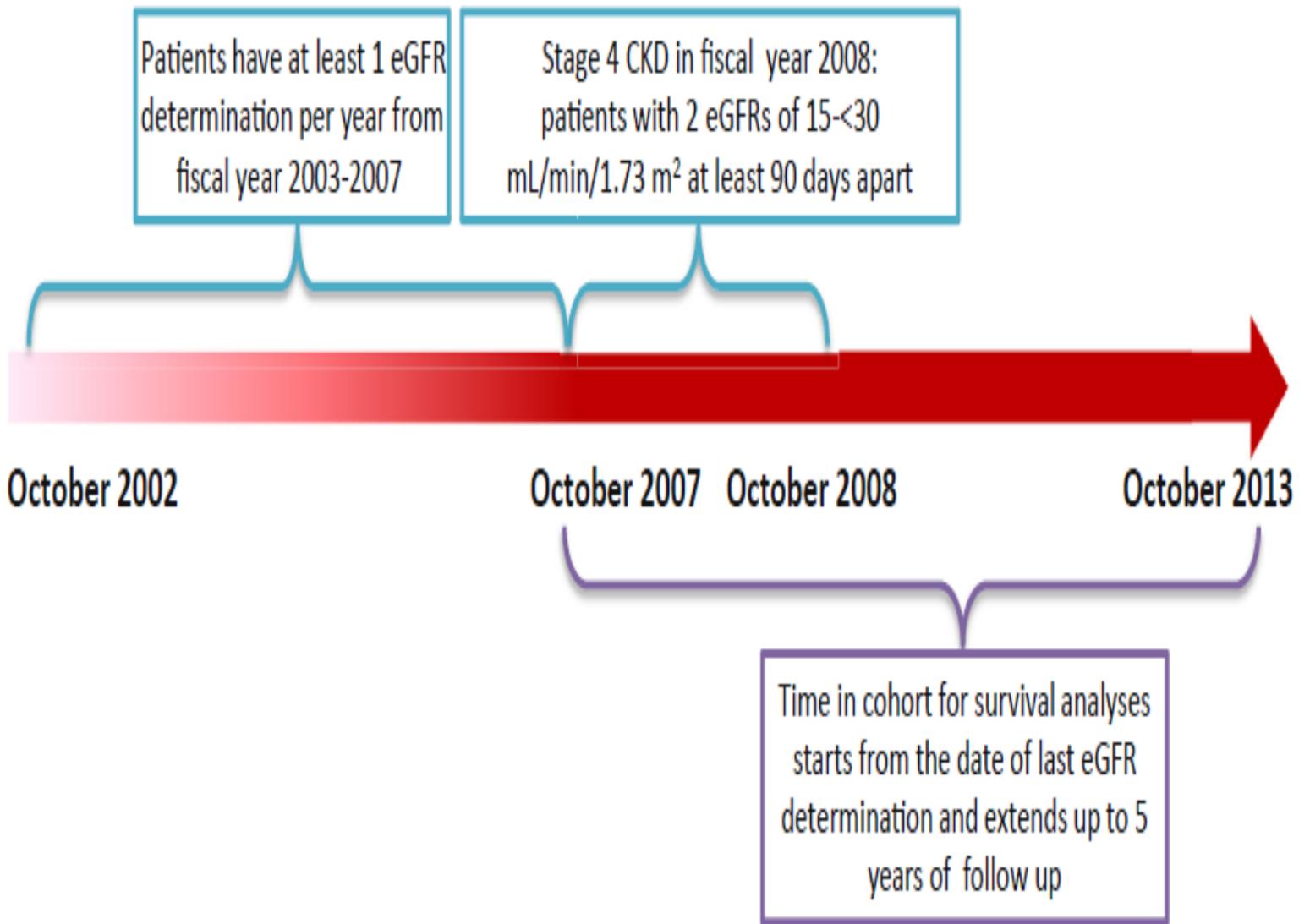


Figure 1. Timeline of cohort assembly. Abbreviations: CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate.



Figure 2. Functional characterization of kidney function trajectories yields 3 trajectory classes: trajectory class 1, consistent slow decline; trajectory class 2, consistent fast decline; and trajectory class 3, early nondecline and late fast decline. Abbreviation: eGFR, estimated glomerular filtration rate.

Table 6. Kidney Function Trajectories and Risk for Death Within 1, 2, 3, 4, and 5 Years

Risk for Death Within:	Trajectory Class 2	Trajectory Class 3
1 y	1.17 (1.10-1.28) ^a	1.29 (1.18-1.42) ^a
2 y	1.13 (1.06-1.21) ^a	1.19 (1.11-1.29) ^a
3 y	1.13 (1.07-1.19) ^a	1.12 (1.05-1.20) ^a
4 y	1.12 (1.07-1.18) ^a	1.06 (1.00-1.13)
5 y	1.11 (1.06-1.16) ^a	1.04 (0.98-1.10)

Note: Values are given as hazard ratio (95% confidence interval). Trajectory class 1 (consistent slow decline) is referent group with all hazard ratios of 1.00. Trajectory class 2, consistent fast decline; trajectory class 3, early nondecline and late fast decline. Models adjusted for receipt of nephrology care, number of hospitalizations, age, sex, race, qualifying estimated glomerular filtration rate, cerebrovascular disease, cardiovascular disease, dementia, diabetes mellitus, hepatitis C virus, human immunodeficiency virus, hypertension, hyperlipidemia, chronic lung disease, and peripheral artery disease.

^aStatistically significant.

Session Outline

- Overview of laboratory data in the VA Managerial Cost Accounting (MCA) National Data Extracts (NDEs)
- Our experience using VA MCA NDE laboratory data
- Overview of laboratory data in the VA Corporate Data Warehouse (CDW)
- Our experience using VA CDW laboratory data

Poll Question #3: Have you worked with lab data in CDW?

- Yes
- No

Lab data in CDW

- CDW is a parallel and increasingly relevant source of lab data in VA
- Both MCA and CDW lab data are derived from Vista at individual medical centers
- CDW includes all lab tests, not select group of tests (as in MCA)
- Available back to FY 2000

Labs in CDW

- Format
 - SQL
- Structure
 - LabChem domain with multiple tables
 - Suggestion is to use PatientLabChem for results
- Identifies for type of lab test
 - LOINCSID (links to LOINC using dim.LOINC metadata)
 - LabChemTestSID (links to LabChemTestName using Dim.LabChemTest metadata)

Identifying labs in CDW

- We usually start with LOINC
 - In our experience with the labs we have done this for, LOINC usually is there in 85-90% of the data we want (early cohorts)
- Text search with lab names
 - For any given lab you can end up with 100s of names
 - Will need to decide on the best approach, may vary by lab
 - We have in the past done a few key search terms, then focused on the names that occurred the most

Issues

- Unexpected lab chem test names
 - Names aren't always straightforward, hard to tell if they are valid
- Result contamination
 - Related measures with what should be different LOINC codes may find their way into each other's results

Session Outline

- Overview of laboratory data in the VA Managerial Cost Accounting (MCA) National Data Extracts (NDEs)
- Our experience using VA MCA NDE laboratory data
- Overview of laboratory data in the VA Corporate Data Warehouse (CDW)
- Our experience using VA CDW laboratory data

Examples of laboratory results we have obtained from CDW

- Monocyte count and monocyte per leukocyte percent
- Erythrocyte sedimentation rate (ESR)
- C-reactive protein (CRP)

Manual count of monocytes with units

- Some units of data from a query of CDW PatientLabChem for a LOINC of 742-7, which corresponds to “Monocytes [# /volume] in Blood by Automated count”
- Suggests some of this data is probably not monocyte count

Unit	%	Mean (SD)
%	2.67	8.08 (5.56)
%	8.14	4.40 (3.31)
k/cmm	22.04	0.64 (0.33)
k/cumm	18.79	0.66 (0.34)
k/uL	5.58	0.69 (0.76)
k/uL	9.98	0.61 (0.50)
X10 ⁻³ /uL	4.06	0.66 (0.33)

Monocyte lab name

- Query of CDW PatientLabChem data by LabChemTestName, via the LabChemTest metadata, for lab chem test names that have “MONOCYTE” in it. Returns 260 possible names
- Is it worth it to clean the names that occur infrequently if the majority of the data happens with just of a few of the names?

Lab Chem Test Name	%
.Monocytes(Absolute)Dc 'd 2/5/08	0.00
Zcd14/monocyte	0.00
Monocyte %	15.27
Monocytes (Manual)	0.11

Lessons Learned

- Related lab measures often find their way into each other
- Important to consider lab value ranges and units, not just the identifier
- It is important to consider the context of when/why the lab value was taken, as this may result in bias
- Relationships between lab values and outcomes are not always linear

Summary Points

- There are several viable options for obtaining VA-wide lab data
- CDW lab data overcome many of the limitations of the MCA LAR file
- Lab data from CDW require more manipulation by end-user to transform into a usable format compared with MCA

VIReC Intranet Resources

- VIReC Research User Guides (RUGs)
 - <http://vaww.virec.research.va.gov/RUGs/RUGs-Index.htm> (VA Intranet only)
- VHA Corporate Data Warehouse (CDW)
 - <http://vaww.virec.research.va.gov/CDW/Overview.htm> (VA Intranet only)
- CDW SharePoint Site & VINCI SharePoint Site
 - Send va.gov email to VIReC@va.gov for URL

VHA Data Portal

- MCA (formerly DSS) National Data Extracts (NDEs)
 - [http://vaww.vhadataportal.med.va.gov/DataSources/MCA\(formerlyDSS\)NDEs.aspx](http://vaww.vhadataportal.med.va.gov/DataSources/MCA(formerlyDSS)NDEs.aspx) (VA Intranet only)
- VHA Corporate Data Warehouse (CDW)
 - <http://vaww.vhadataportal.med.va.gov/DataSources/CDW.aspx> (VA Intranet only)

VIReC Help

- HSRData Listserv
 - Join at the VIReC website
 - Discussion among data stewards, managers, and users
 - Past messages in archive (on VA Intranet)
- VIReC Help Desk
 - VIReC staff will answer your question and/or direct you to available resources on topics
 - VIReC@va.gov; 708-202-2413

For more information

- Clinical Epidemiology Center
 - stlcec@va.gov
 - Benjamin.Bowe@va.gov
 - Yan.Xie3@va.gov

