Measuring Laboratory Use and Results Using the VA DSS
National Lab Data

April 7, 2014

Presented by:
Ann O’Hare, MA MD
Adam Batten, BA
Jeff Todd-Stenberg, BA
Daniel Bertenthal, MPH
Session Outline

- Overview of laboratory data in the VA Decision Support System (DSS)
- Our experience using VA DSS laboratory data
- Overview of laboratory data in the VA Corporate Data Warehouse (CDW)
- Our experience using VA CDW laboratory data
Audience Poll

- Have you worked with lab data in DSS?
  - Yes
  - No

- Have you worked with serum creatinine data in DSS?
  - Yes
  - No
**DSS Overview**

- **What is DSS?**
  - VA’s managerial cost accounting and executive information system

- **What is its primary purpose?**
  - To provide managerially-useful information (e.g., productivity measures, costs per unit of work, quality assessment) to:
    - Managers
    - Undersecretary for Health
    - Secretary
    - Congress
DSS Source Data

Financial Systems
- Payroll
- Building Depreciation
- etc.

Workload
- VistA Packages
  - Laboratory
  - Pharmacy
  - Nursing
  - etc.

Patient Info
- VistA (CPRS)
- NPCD
- PTF
- etc.

DSS Processes

DSS Data
DSS National Data Extracts (NDE)

DSS Processes
- Financial Systems
- Workload
- Patient Info

DSS Data

National Data Extracts
- LAB
- OUT
- RAI
- PHA
- LAR
- DIS
National Data Extracts (NDEs)

**Clinical NDEs**

- **LAB**
  - Workload and costs
  - Test-level records

- **LAR**
  - Laboratory results for a defined list of tests (currently 91)
  - Test-level records
DSS NDEs

- Schedule: Monthly or quarterly
- Cumulative year-to-date
- Lab data from FY2000 (LAR) or 2002 (LAB)
LAR TESTS used in our work

- Serum creatinine—DSSLARNO 31
- Urine albumin-to-creatinine ratio—DSSLARNO 56
- Hepatitis C antibody –DSSLARNO 89
DSS NDE Data Formats

- **Reports, Data Cubes**
  - VHA Support Service Center (VSSC)

- **SAS Datasets**
  - Separate LAR dataset for each VISN and each Fiscal year
  - Discontinued after FY 2012
  - Removed from Austin Information Technology Center (AITC) mainframe in March 2013
  - LAR SAS datasets for 2000-2012 now reside in VINCI

- **SQL tables**
  - Single SQL file for all LAR results, across fiscal years and VISNS
  - Available from 2005 onward
  - Reside in Corporate Data Warehouse (CDW)
Prevalence of low estimated GFR (eGFR) by age group.

O’Hare et al. JASN 2006;17:846-853
Baseline eGFR threshold below which risk for ESRD exceeded risk for death for each age group.

O'Hare et al. JASN 2007;18:2758-2765
White/Black Racial Differences in Risk of End-stage Renal Disease and Death

Choi et al. The American Journal of Medicine, Volume 122, Issue 7, 2009, 672 - 678
From: **Association of Hepatitis C Seropositivity With Increased Risk for Developing End-stage Renal Disease**

Table 2. End-stage Renal Disease (ESRD) Incidence Rates and Hazards Ratios (HR) by Hepatitis C Virus (HCV) Seropositivity

<table>
<thead>
<tr>
<th>Variable</th>
<th>HCV Antibody</th>
<th>Observation Person-Years</th>
<th>ESRD Events</th>
<th>Unadjusted Rate per 1000 Person-Years (95% CI)</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>4393</td>
<td>3.05 (2.96-3.14)</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td>760</td>
<td>4.26 (3.97-4.57)</td>
<td>1.40 (1.29-1.51)</td>
</tr>
<tr>
<td>Stratified by age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-49</td>
<td>Negative</td>
<td>340 370</td>
<td>425</td>
<td>1.25 (1.14-1.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>83 409</td>
<td>243</td>
<td>2.91 (2.57-3.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1087</td>
<td>2.37 (2.23-2.51)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>333</td>
<td>4.79 (4.30-5.33)</td>
<td></td>
</tr>
<tr>
<td>Stratified by age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59</td>
<td>Negative</td>
<td>458 725</td>
<td>1244</td>
<td>3.87 (3.66-4.09)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>69 589</td>
<td>113</td>
<td>8.04 (6.69-9.67)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1637</td>
<td>5.08 (4.83-5.33)</td>
<td></td>
</tr>
<tr>
<td>Stratified by age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>Negative</td>
<td>321 192</td>
<td>71</td>
<td>6.23 (4.93-7.86)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>14 047</td>
<td>426</td>
<td>1.62 (1.46-1.79)</td>
<td></td>
</tr>
<tr>
<td>≥70</td>
<td>Negative</td>
<td>322 539</td>
<td>71</td>
<td>1.22 (0.96-1.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>11 403</td>
<td>426</td>
<td>1.08 (0.85-1.37)</td>
<td></td>
</tr>
<tr>
<td>Stratified by eGFR, mL/min per 1.73 m²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥60</td>
<td>Negative</td>
<td>1 219 523</td>
<td>308</td>
<td>0.25 (0.23-0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>163 846</td>
<td>106</td>
<td>0.65 (0.54-0.78)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1039</td>
<td>5.01 (4.71-5.32)</td>
<td></td>
</tr>
<tr>
<td>Stratified by age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-59</td>
<td>Negative</td>
<td>207 484</td>
<td>228</td>
<td>17.15 (15.06-19.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>13 295</td>
<td>192.55 (185.83-199.51)</td>
<td>1.62 (1.46-1.79)</td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>Negative</td>
<td>15 819</td>
<td>228</td>
<td>3.45 (2.99-3.98)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>1307</td>
<td>326.03 (296.49-358.50)</td>
<td>1.28 (1.15-1.44)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; eGFR, estimated glomerular filtration rate.
*Adjusted for all covariates (age, sex, race/ethnicity, diabetes mellitus, hypertension, human immunodeficiency virus infection, congestive heart failure, coronary heart disease, peripheral vascular disease, chronic obstructive pulmonary disease, cerebrovascular disease, substance abuse, and baseline eGFR), minus the covariate on which stratified.
Primary disease causing ESRD by HIV/diabetes (DM) status among white (A) and black (B) patients.

**A**

<table>
<thead>
<tr>
<th>HIV/DM Category</th>
<th>Percent of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>-HIV/-DM (n=3,991)</td>
<td></td>
</tr>
<tr>
<td>+HIV/-DM (n=13)</td>
<td></td>
</tr>
<tr>
<td>+DM/-HIV (n=6,749)</td>
<td></td>
</tr>
<tr>
<td>+HIV/+DM (n=16)</td>
<td></td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>HIV/DM Category</th>
<th>Percent of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>-HIV/-DM (n=1,425)</td>
<td></td>
</tr>
<tr>
<td>+HIV/-DM (n=129)</td>
<td></td>
</tr>
<tr>
<td>+DM/-HIV (n=2,775)</td>
<td></td>
</tr>
<tr>
<td>+HIV/+DM (n=50)</td>
<td></td>
</tr>
</tbody>
</table>

Choi A I et al. JASN 2007;18:2968-2974

©2007 by American Society of Nephrology
From: *Prognostic Implications of the Urinary Albumin to Creatinine Ratio in Veterans of Different Ages With Diabetes*

Trajectories of Kidney Function Decline in the 2 Years Before Initiation of Long-term Dialysis

Strengths and limitations of DSS Laboratory data

- Availability of laboratory measures allowed for major advance beyond work based on diagnostic codes
- Easy for end-user (single DSSLARNO for each test)
- Incomplete capture of all relevant lab tests at all medical centers, especially for earlier years
- Lack of transparency about mapping process ("black box"), especially for earlier years
LOINC Codes
http://www.loinc.org

- LOINC: Logical Observation Identifier Names and Codes
- Highly specific - Identifies test, method of analysis, specimen source
- Lab results (LAR) for DSS records pulled based on LOINC, implemented nationwide back to FY 2009 onward, available in LAB from FY 2013 onward
- Results in better match between LAB and LAR records compared with previous method (based on test names)
- VistA LOINC file and DSS LOINC file have not always contained the same version of the LOINC code set.
Audience Poll

Have you worked with lab data in CDW?
- Yes
- No

Have you worked with serum creatinine or proteinuria data in CDW?
- Yes
- No
Lab data at CDW

• CDW is a parallel and increasingly relevant source of lab data in VA
• Both DSS and CDW lab data are derived from VistA at individual medical centers
• CDW includes all lab tests, not select group of tests (as in DSS)
• Disadvantage (or advantage) is that it follows VistA test name format at each medical center
• Available back to FY 2000
• Contains LOINC code
From: Interpreting Treatment Effects From Clinical Trials in the Context of Real-World Risk Information: End-Stage Renal Disease Prevention in Older Adults

Lab Data Mapping

Query Lab Table

Update Query Criteria

Query Lab Tests/Results Analyzed

Present Curious Results to Clinician (Ann)

Remove Obvious Lab Test/Results
<table>
<thead>
<tr>
<th>LabChemTestName</th>
<th>FREQ</th>
<th>Missing Numeric</th>
<th>min</th>
<th>mean</th>
<th>median</th>
<th>IQR</th>
<th>p90</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>URINE PROTEIN</td>
<td>12114830</td>
<td>10565125</td>
<td>0</td>
<td>64.6122</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>38000</td>
</tr>
<tr>
<td>BUN/CREATININE RATIO</td>
<td>1416565</td>
<td>2169</td>
<td>-217</td>
<td>16.9531</td>
<td>16</td>
<td>12.4</td>
<td>25</td>
<td>310</td>
</tr>
<tr>
<td>MICROALBUMIN</td>
<td>761995</td>
<td>156860</td>
<td>0</td>
<td>54.0178</td>
<td>7.4</td>
<td>2</td>
<td>8</td>
<td>64437</td>
</tr>
<tr>
<td>UREA NITROGEN/CREATININE RATIO</td>
<td>734505</td>
<td>276</td>
<td>-37.5</td>
<td>15.6418</td>
<td>14.6</td>
<td>11.3</td>
<td>23.3</td>
<td>430</td>
</tr>
<tr>
<td>URINE PROTEIN (UA)</td>
<td>503761</td>
<td>449626</td>
<td>0</td>
<td>77.8838</td>
<td>30</td>
<td>30</td>
<td>150</td>
<td>500</td>
</tr>
<tr>
<td>PROTEIN, URINE</td>
<td>366745</td>
<td>328964</td>
<td>0</td>
<td>85.0025</td>
<td>30</td>
<td>30</td>
<td>133</td>
<td>35650</td>
</tr>
<tr>
<td>MICROALBUMIN/CREATININE RATIO</td>
<td>363984</td>
<td>58443</td>
<td>-8400</td>
<td>89.8153</td>
<td>13.6</td>
<td>5.2</td>
<td>146.5</td>
<td>733000</td>
</tr>
<tr>
<td>URINE PROTEIN (V2)</td>
<td>355594</td>
<td>290275</td>
<td>0.2</td>
<td>73.4863</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td>URINE CREATININE</td>
<td>343147</td>
<td>15910</td>
<td>-0.2</td>
<td>121.547</td>
<td>108</td>
<td>68.4</td>
<td>214.1</td>
<td>98317.5</td>
</tr>
<tr>
<td>URINE PROTEIN (Dipstick)</td>
<td>244514</td>
<td>227082</td>
<td>1</td>
<td>1.4692</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>URINE PROTEIN (URINALYSIS)</td>
<td>243078</td>
<td>196736</td>
<td>0.1</td>
<td>61.9557</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>2000</td>
</tr>
<tr>
<td>CREATININE (ORDER URINE SEPARATELY)</td>
<td>235703</td>
<td>2549</td>
<td>0.1</td>
<td>4.3256</td>
<td>1.2</td>
<td>1</td>
<td>2.4</td>
<td>957</td>
</tr>
<tr>
<td>Microalbumin</td>
<td>222681</td>
<td>38259</td>
<td>0</td>
<td>45.8758</td>
<td>7.9</td>
<td>2.5</td>
<td>77.1</td>
<td>63333.3</td>
</tr>
<tr>
<td>PROTEIN, URINE</td>
<td>211453</td>
<td>169779</td>
<td>0</td>
<td>72.1782</td>
<td>30</td>
<td>15</td>
<td>156</td>
<td>6100</td>
</tr>
<tr>
<td>MICROALBUMIN/CREAT RATIO</td>
<td>177510</td>
<td>30556</td>
<td>0</td>
<td>67.302</td>
<td>10.3</td>
<td>4</td>
<td>123</td>
<td>42515</td>
</tr>
<tr>
<td>CREATININE, URINE</td>
<td>175965</td>
<td>8718</td>
<td>0</td>
<td>122.532</td>
<td>109.5</td>
<td>68.3</td>
<td>217.61</td>
<td>4446</td>
</tr>
<tr>
<td>Urine Protein</td>
<td>166808</td>
<td>114394</td>
<td>0</td>
<td>48.1698</td>
<td>20</td>
<td>10</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Protein urine</td>
<td>156337</td>
<td>123408</td>
<td>0</td>
<td>63.2836</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>CREATININE, URINE</td>
<td>153673</td>
<td>8712</td>
<td>0</td>
<td>135.703</td>
<td>109</td>
<td>67.9</td>
<td>218.2</td>
<td>5105</td>
</tr>
<tr>
<td>MICROALBUMIN (V2)</td>
<td>147825</td>
<td>18076</td>
<td>0</td>
<td>7.6405</td>
<td>2.24</td>
<td>1</td>
<td>17.79</td>
<td>2297</td>
</tr>
<tr>
<td>ALBUMIN/CREATININE RATIO</td>
<td>146467</td>
<td>40865</td>
<td>0</td>
<td>118.549</td>
<td>11.7</td>
<td>6</td>
<td>67.2</td>
<td>3166667</td>
</tr>
<tr>
<td>MICROALBUMIN/CREATININE RATIO (V2)</td>
<td>144532</td>
<td>17108</td>
<td>0</td>
<td>71.0015</td>
<td>18</td>
<td>8</td>
<td>163.1</td>
<td>10948</td>
</tr>
<tr>
<td>MICROALBUMIN (mg/dL)</td>
<td>143812</td>
<td>13884</td>
<td>0</td>
<td>13.1453</td>
<td>1.9</td>
<td>0.7</td>
<td>24.2</td>
<td>5551.8</td>
</tr>
<tr>
<td>PROTEIN, URINE (STICK)</td>
<td>143635</td>
<td>109156</td>
<td>0</td>
<td>79.1281</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>URINE PROTEIN (QUAL)</td>
<td>131929</td>
<td>107564</td>
<td>2</td>
<td>84.795</td>
<td>30</td>
<td>30</td>
<td>300</td>
<td>2000</td>
</tr>
<tr>
<td>PROTEIN - (dipstick urine)</td>
<td>121742</td>
<td>119768</td>
<td>0.2</td>
<td>63.188</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td>Urine MICROALBUMIN</td>
<td>117278</td>
<td>8599</td>
<td>0.1</td>
<td>15.4437</td>
<td>2.1</td>
<td>0.8</td>
<td>28.7</td>
<td>7582</td>
</tr>
<tr>
<td>MICROALBUMIN (Spur Ur)</td>
<td>111538</td>
<td>27175</td>
<td>0.2</td>
<td>3.7238</td>
<td>1.4</td>
<td>0.6</td>
<td>10.3</td>
<td>133</td>
</tr>
<tr>
<td>Urine PROTEIN-QUAL-0</td>
<td>108679</td>
<td>85334</td>
<td>30</td>
<td>77.6783</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>300</td>
</tr>
<tr>
<td>RATIO OF MICROALBUMIN/CREATININE</td>
<td>104621</td>
<td>24589</td>
<td>0</td>
<td>36.325</td>
<td>11</td>
<td>5</td>
<td>93</td>
<td>3445</td>
</tr>
<tr>
<td>URINE DAU CREATININE</td>
<td>96291</td>
<td>1856</td>
<td>0</td>
<td>143.152</td>
<td>131.4</td>
<td>82</td>
<td>255.7</td>
<td>2860.1</td>
</tr>
<tr>
<td>Urine PROTEIN 2</td>
<td>94523</td>
<td>2814</td>
<td>34</td>
<td>161725</td>
<td>3.7</td>
<td>0.5</td>
<td>21.5</td>
<td>2.05</td>
</tr>
</tbody>
</table>
There are many spellings for the same test value. Fortunately, the bulk of the tests have similar names. Handful of oddities .neg, +- , <1, etc... Mysterious values taken back to Ann for review.
Audience Poll

Experience level with regular expressions?
- Never heard of them
- Somewhat familiar
- Use them occasionally
- This is your first thought in the morning
  #^[><](\d+\.?\d?)#
DATA step1;
if _N_ = 1 then do;
  re1 = prxparse("#1\(d\)\+$"); /*3+*/
  re2 = prxparse("#1\(<\(d+\|d?d\)#\)\+$"); /*10+*/
  re3 = prxparse("#1\(d\)+\/[MOGNAD\(d\)#\)\+$"); /*300MG/DL*/
  re4 = prxparse("#1\(d\)+-(\+\)\+$"); /*100-200*/
  re5 = prxparse("#1\(<\(d+\|d?d\)#\)\+$"); /*<25*/
  re6 = prxparse("#1\(d\)+\+$"); /*+1*/
  re7 = prxparse("#1\(RAT\)\+$"); /*TRACE*/
  re8 = prxparse("#1\(POS\)\+$"); /*POSITIVE*/
  re9 = prxparse("#1\(NEG\)\+$"); /*NEGATIVE*/
  re92 = prxparse("#1\(E|^GR\)\+$"); /*NEG*/
  re10 = prxparse("#1\(\d{2}\)\+$"); /*+(20*/
  re11 = prxparse("#1\(d\)\+$"); /*007*/
  re12 = prxparse("#1\(d\)\+$"); /*1000*/
  re13 = prxparse("#1\(ERATE\)\+$");
  re14 = prxparse("#1\(ALL\)\+$");
  re15 = prxparse("#1\(RGE\)\+$");
  re16 = prxparse("#1\(d\)\+$"); /*30*/
end;
retain re1 re2 re3 re4 re5 re6 re7 re2 re9 re92 re10 re11 re12 re13 re14 re15 re16;
length one two eight upcode $2;
set dp.Ppckd_cdw.labchemtestresults;
where trim(left(upcase(LabChemTestName))) like '%URINE%'
  AND trim(left(upcase(LabChemTestName))) like '%PROTEIN%'
  AND trim(left(upcase(LabChemTestName))) like '%24%'
  AND trim(left(upcase(LabChemTestName))) not like '%12%'
  AND trim(left(upcase(LabChemTestName))) not like '%ELECTRO%'
  AND trim(left(upcase(LabChemTestName))) not like '%EP%'
  AND trim(left(upcase(LabChemTestName))) not like '%LC%';
upcode="NA";
result=compress(left(upcase(LabChemResultValue)));
RESULT=result;
Lab Data Mapping

- Input from clinician is invaluable
- Most text data is easy to map
- Programmers- master REGEXP, you won’t be sorry.
- Future directions? Compare use of LOINC to text mapping? Index every possible dipstick lab test? ...
Summary points

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

DSS Research User Guide
- http://www.virec.research.va.gov/RUGs/RUGs-Index.htm

VHA Corporate Data Warehouse (CDW)
VIReC Help

HSRData Listserv
- Join at the VIReC website
- Discussion among >650 data stewards, managers, and users
- Past messages in archive (on 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
VA Intranet Websites

- **CDW SharePoint Site**
  - Send va.gov email to VIReC@va.gov for URL

- **VINCI SharePoint Site**
  - Send va.gov email to VIReC@va.gov for URL
VIReC@va.gov
Upcoming Seminar

May 5, 2014

Assessing Race and Ethnicity

Maria Mor, PhD