

Working with the CDW Health Factors Domain: Tobacco Use Status

Paul G. Barnett, PhD^{1,2} Adam Chow^{1,2}

Nicole E. Flores^{1,2} Sonia Duffy, PhD³

1. Health Economics Resource Center

2. Center for Innovation to Implementation, Menlo Park, CA

3. Center for Clinical Management Research, Ann Arbor, MI

September 15, 2015

Acknowledgements

- HSR&D QUERI Rapid Response Project (RRP 11-266)
 - Study advisors: Suvarna Akki, Sara Bailey, Jean Beckham, Timothy Carmody, Kim W. Hamlett-Berry, Miles E. McFall & Andrew J. Saxon , Elizabeth Gifford, James Halloran, Ware Kuschner & Richard H. Pham
-

Overview of Presentation

1. Corporate Data Warehouse (CDW)
health factors data
 2. Data analysis and evaluation
 3. Changes in Veteran tobacco use
 4. Ongoing work
 5. Conclusions
-

CDW health factors data

- Entries generated by clinical reminders and other VISTA scripts
 - Patient, date, provider, and 40 characters representing responses to prompts
 - 1,071 unique text strings with “smoke” or “tobacco”
 - 478 relevant to tobacco use status
-

Example health factor entries

Health Factor Type

LIFETIME NON-TOBACCO USER

LIFETIME NON-SMOKER

CURRENT SMOKER

PREVIOUS SMOKER

FORMER SMOKER - <100 LIFETIME CIGARETTES

TOBACCO

CURRENT NON-SMOKER

CURRENT TOBACCO USER

LIFETIME NON-USER OF TOBACCO

QUIT TOBACCO IN THE LAST 12 MONTHS

QUIT TOBACCO >7 YEARS AGO

QUIT TOBACCO >12 MO & <7 YRS AGO

Methods

- We classified text entries
 - Current user
 - Quit within 1 year
 - Quit 1-7 years
 - Quit more than 7 years
 - Never used tobacco
 - Categories needed to implement tobacco use screening guidelines
-

Methods

- Often more than one entry per person
 - We relied on the most specific entry
 - E.g. ambiguous entry “Former smoker” resolved by accompanying entry “Quit <1 year”
 - Small number of entries could not be classified
-

CDW health factors evaluation

- 14.43 million tobacco use assessments of 4.96 million patients 2009-2011
 - Among 5.7 million VHA users of VHA care in FY2011
 - 70.3% had a timely assessment of tobacco use status recorded in health factors data
 - Facility-level completeness varied from 26.4% to 90.6%
-

Changes in Veteran tobacco use

Follow-up of current tobacco users

- We identified 1.0 million current users of tobacco in FY2009
- Within 24 months:
 - 74.8% had a follow-up tobacco use assessment
 - 4.3% had died

Predictors of tobacco cessation

- Among those with follow-up assessment
 - 119,321 (15.8%) of 754,504 no longer using tobacco
- Diagnosis codes used to identify illnesses of at baseline
- Multivariate logistic regression age, gender, region, and chronic illnesses

Predictors of tobacco cessation

Chronic Illness	Odds Ratio
Age > 70	1.25
Asthma	1.36
Cancer	1.18
Lung Cancer	1.38
Heart failure	1.26
COPD	0.91
Pneumonia	1.28
Schizophrenia	0.74
Alcohol abuse or dependence	0.76
Drug abuse or dependence	0.85

Partial list of parameters from multivariate logistic regression; all with $p < 0.001$

Follow-up of recent quitters

- We identified 73 thousand recent quitters (quit < 1 year) in FY 2009
- Within 24 months
 - 71.7% had follow-up assessment
 - 5.8% died
- Among recent quitters with follow-up
 - 38.2% had relapsed (resumed tobacco use)

Predictors of tobacco relapse

- 413,979 former tobacco users with follow-up data
- Relapsed by 38,932 (9.4%)
 - 38.2% relapse in those quit < 1 year
 - 13.0% relapse in those quit 1-7 years
 - 2.7% relapse in those quit > 7 years

Predictors of tobacco relapse

Chronic Illness	Odds Ratio
Quit < 1 year	15.42
Quit 1 – 7 years	3.98
Age > 70	0.34
Asthma	0.89
Lung Cancer	0.91
COPD	1.14
Pneumonia	1.16
Schizophrenia	1.42
Alcohol abuse or dependence	1.31
Drug abuse or dependence	1.60

Partial list of parameters from multivariate logistic regression; all with $p < 0.001$

Tobacco cessation and causality

- Poor health may precede tobacco cessation
- Care must be taken in making inferences about the effects of quitting
 - e.g., Studies have found that health care costs increase immediately after quitting

Ongoing work

- Use of dataset for long-term follow-up in a trial of tobacco cessation in VA residential care (Liz Gifford, PI)
- Evaluation of effectiveness of tobacco pharmacotherapy (Sonia Duffy, PI)

Conclusions

- Health factors data require interpretation
 - See HERC technical report #28
 - www.herc.research.va.gov
 - Data can be useful
 - Evaluation of cessation programs
 - Tobacco status as a covariate
 - Affordable Care Act mandates field for tobacco in Electronic Health Record
-

Identifying, Cleaning, and Validating Functional Status Data in the CDW Health Factors Domain

by Rebecca Brown, MD, MPH



Objectives

- Introduction to CDW Health Factors (HFs) Domain
- Case Study: identifying and cleaning HFs data on functional status
 - Lessons learned – opportunities and challenges
- Ongoing work: validating CDW HFs 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)

Health Factors Domain: Overview

- VistA data elements
- Often used to capture results of “clinical reminders”
- Clinical reminders
 - Automated alerts that trigger providers to perform evidence-based tests, other interventions
 - E.g.: measure HbA1C, screen for smoking, screen for colon cancer
- Not standardized across VA

Health Factors Domain: Organization

- Like other CDW data, Dimension and Fact Tables
- Tables contain name of HF plus linked data
 - Station where collected
 - Date and time when collected
 - Type of patient encounter when collected
 - Patient ID
 - Etc.

Health Factors Domain: Examples

History Data			
9999999.64	1	NON-TOBACCO USER	6~TOBACCO
9999999.64	2	CURRENT SMOKER	6~TOBACCO
9999999.64	3	CURRENT NON-SMOKER	6~TOBACCO
9999999.64	4	PREVIOUS SMOKER	6~TOBACCO
9999999.64	5	LIFETIME NON-SMOKER	6~TOBACCO
9999999.64	6	TOBACCO	6~TOBACCO
9999999.64	7	TB STATUS	7~TB STATUS
9999999.64	8	TB - TX COMPLETE	7~TB STATUS
9999999.64	9	TB - TX INCOMPLETE	7~TB STATUS
9999999.64	10	TB - TX UNKNOWN	7~TB STATUS
9999999.64	11	TB - TX UNTREATED	7~TB STATUS
9999999.64	12	ALCOHOL USE	~
9999999.64	13	HISTORY OF AN ALCOHOL PROBLEM	12~ALCOHOL USE
9999999.64	14	DRIVING UNDER THE INFLUENCE	12~ALCOHOL USE
9999999.64	15	FAMILY HX OF ALCOHOL ABUSE	12~ALCOHOL USE
9999999.64	16	PREV. SCREEN ETOH PROBLEM	12~ALCOHOL USE
9999999.64	17	HEAVY DRINKER (3 OR MORE/DAY)	12~ALCOHOL USE
9999999.64	18	DRINKING ALONE	12~ALCOHOL USE
9999999.64	19	BINGE DRINKING	12~ALCOHOL USE
9999999.64	20	HX BREAST CANCER	21~BREAST CANCER
9999999.64	21	BREAST CANCER	21~BREAST CANCER
9999999.64	22	FAMILY HX BREAST CANCER	21~BREAST CANCER
9999999.64	23	PREV. BREAST CANCER SCREENING	21~BREAST CANCER

Poll Question #2: I have previously used data from the CDW Health Factors domain.

- Yes
- No

Health Factors: Case Study

- Study: validate HFs measures of functional status
 - Functional status = ability to perform activities of daily living (ADLs: e.g., bathing, dressing)
 - ~6 years ago, VA GEC encouraged medical centers to collect these data using clinical reminders (once/yr, aged 75+, primary care)

Health Factors: Case Study

- Unknown how many centers collecting data or if data are valid/encoded accurately
- Objective: validate functional status data collected in older adults attending primary care appointments

Identifying functional status data

- First step: identify, extract, and clean measures of functional status in Health Factors domain
 - Extracted HFs data for people aged 65+ in 2009-2013 = 238,150 unique HFs at 129 stations

Favorite HFs

1. CCHT UNCERTAIN WHAT TO DO=NEARLY ALWAYS
2. WHAT MAKES PAIN BETTER:CAR RIDING
3. FEEDING THE TIGER
4. URINARY INCONTINENCE-DIRECT OBSERVATION
5. COUNT SILVERWARE BEFORE/AFTER MEALS
6. TAKE YOUR MIND FOR A WALK
7. MOVING THROUGH THE SWAMP
8. PATIENT REFUSED

Identifying functional status data

- Developed systematic approach to identify
 - Searched list of HFs for key words related to daily activities (e.g., bath, dress, transfer, medic)
 - After key word search, manual search to identify misspellings, etc.
 - → 2178 HFs from 51 stations

Example HFs

FUNCTIONAL SCREEN BATHING

FUNCTIONAL SCREEN DRESSING

FUNCTIONAL SCREEN EATING

FUNCTIONAL SCREEN FOR ADLs

FUNCTIONAL SCREEN MOBILITY/TRANSFERRING

FUNCTIONAL SCREEN TOILETING

Identifying functional status data

- Narrowed list using following eligibility criteria:
 - Appropriate coding: 2-level variable ADL/IADL (e.g., bathing independent/bathing dependent)
 - Complete: Includes 5 ADLs and 8 IADLs
 - Clinically plausible: ~10-20% of patients at given station coded as “dependent”
 - Outpatient setting: primary care encounter code
 - → 442 HFs from 17 stations

Example - eligible HFs

BATHING INDEPENDENT

BATHING NOT INDEPENDENT

DRESSING INDEPENDENT

DRESSING NOT INDEPENDENT

FEEDING INDEPENDENT

FEEDING NOT INDEPENDENT

TOILETING INDEPENDENT

TOILETING NOT INDEPENDENT

TRANSFERRING INDEPENDENT

TRANSFERRING NOT INDEPENDENT

Challenges

- Many stations collecting data – but most not usable
 - Many HF labels don't capture categorical data – e.g., single label: “ADL screen completed”
 - Missing items – e.g., only 4 of 5 key ADLs
 - Clinically implausible – 100% of patients coded as needing help with 1+ ADLs

Challenges

- No standard instrument used to assess function
 - Contacted eligible sites to get CPRS screen shots
 - Variety of instruments in use → narrowed eligible sites from 17 to 15
 - Selected 7 geographically representative sites for validation study
 - From these 7 sites, daily data pulls identify ~150 patients with data collected the previous day

Challenges: Real-time data collection

- Problem: data from one VISN disappeared
- Contacted VISN clinical reminders staff →
 - Network leader directed staff to stop collecting data – too time-consuming
 - Lost data from multiple sites → identify alternative sites, update code for VINCI/locally

Challenge: Real-time data collection

- Problem: no data in daily data pulls
- Contacted VINCI – new HFs changes in CDW
 - Identify new table names/locations, update code for VINCI/locally

Recommendations: using HFs data

- Complexity of HFs (differing instruments, inconsistent uptake, variable data labels) →
- Take time to learn “history” of HF/key issues
 - Leaders involved in implementation, clinical reminders experts, front-line staff

Recommendations: using HFs data

- Algorithms to identify/clean HFs data need checks at multiple levels
 - Ensure consistent instrument (CPRS), encounter type (codes), data categorization, etc.
- Given heterogeneity in HFs, may end up with selected sample – limits generalizability
 - Balance between inclusion/internal consistency

Next steps: validation study

- Completed “internal” checks for logic of HFs
- Now completing “external” checks
 - Compare sensitivity/specificity of CDW HFs measures to “reference standard” of research-collected measures

Ongoing validation study

- Daily: pull list of individuals with relevant HFs collected previous weekday
- Send opt-out letter explaining study (toll-free #)
- If no call within 1 week,
 - Call individuals
 - Assess eligibility
 - Obtain informed consent
 - Administer validated ADL/IADL measures

Ongoing validation study

- ~250 participants enrolled to date
- Anticipate completing data collection next few months

Conclusions

- CDW HFs domain includes unique data with big potential
- ...but requires a lot of upfront work (cleaning, interpretation) to be useful
- Ideally need to validate data: even with use of validated measures, many chances to lose “fidelity”
 - Data entry, data labels, data coding, etc.

Contact

- Rebecca Brown
 - rebecca.brown@va.gov