Good Data Practices Cyberseminar Series 4.0

Focusing on the interaction between research design and data decisions



Health Services Research & Development

VIReC Cyberseminar Archive HSR&D's archive for VIReC-hosted cyberseminars

NOTE: Intranet only seminars are available with VA network access from VIReC's Intranet Cyberseminar Archive » http://vaww.virec.research.va.gov/Cyberseminars/Intranet-Archive.htm

Filter by VIReC series
Search by keyword
Sort columns by clicking on the table headings.

/IReC

Filter by VIReC Series: VIReC Good Data Practices

✓ filter

search

Search VIReC archive:

DATE		VIREC SERIES	PRESENTERS
2015/09/24	Using REDCap: Data management in studies linking primary and secondary data	Good Data Practices	Deppen, Stephen
2015/09/22	Decisions, Decisions, Decisions: Selecting Methods and Tools for Data Analysis	Good Data Practices	Hynes, Denise
2015/09/17	Mind the Gap: Using administrative and claims data to answer your research question	Good Data Practices	Wagner, Todd
2015/09/15	Planning for Data: Early, Often and Ongoing	Good Data Practices	Krein, Sarah
2014/05/29	Reduce, Reuse, Recycle: Planning for Data Sharing	Good Data Practices	Kok, Linda
2014/05/22	Controlled Chaos: Tracking Decisions During an Evolving Analysis	Good Data Practices	Groeneveld, Peter
2014/05/15	<u>"The Living Protocol" – Managing Documentation</u> While Managing Data	Good Data Practices	Maciejewski, Matt
2014/05/08	The Best Laid Plans: Plan Well, Plan Early	Good Data Practices	Garvin, Jennifer
2013/09/13	Research Application.	Good Data Practices	Maciejewski, Matt
2013/09/11	Planning for Data Re-use	Good Data Practices	Kok, Linda
2013/09/10	Managing and Documenting Data Workflow	Good Data Practices	Hynes, Denise
2013/09/09	Early Data Planning for Research	Good Data Practices	Hynes, Denise

Previous Good Data Practices cyberseminar series archived on HSR&D website

http://www.hsrd.research.va.gov/cyberseminars/catalog-archive-virec.cfm?

Poll #1 Good Data Practices Series

Have you attended a Good Data Practice cyberseminar session before the current 2017 series?

Yes

• No



Research Life Cycle



Examples of factors that influence data decisions

- Research question
- Study design
 - Objectives, aims, hypotheses
 - Independent and dependent variables
 - Planned manuscripts
- Available data
- Feasibility testing

Learning objectives of the Good Data Practices 4.0 series

Series' participants will

- Understand how previous research results and conceptual/decision models influence the development of the research question
- Learn how a research question can influence the choice of study design
- Understand ways in which research questions and study designs can affect decisions about data
- Become aware of potential data management and analysis challenges and ways they might be addressed
- Become familiar with potential limitations in VA data sources and examples of ways to address them

Visit our Education page for more information & registration links.



FY '17 Good Data Practice Cyberseminar Series Tuesdays and Thursdays in February, 2017 1:00-2:00 PM (ET)

Date	Торіс	Presenter
Tuesday, February 14, 2017	Incorporating Genomics in Routine Care for Veterans with Colon Cancer: Study Design and Data Decisions	Sara Knight
Thursday, February 16, 2017	Data Use and Data Decisions in a Mixed Methods Study about Hand Hygiene	Heather Reisinger
Tuesday, February 21, 2017	Data Decisions and Quantitative Analysis in a Study Investigating the Impact of Remote ICU Monitoring in VA Hospitals	Mary Vaughan-Sarrazin Amy O'Shea
Thursday, February 23, 2017	Capstone Discussion: The Influence of Research Design on Data Decisions	Discussant: Neil Jordan

Poll #2: Your role as a data user

- What is your role in research and/or quality improvement?
 - a. Research investigator
 - b. Data manager
 - c. Project coordinator
 - d. Clinical Staff
 - e. Operations Staff
 - f. Other (please specify)



Good Data Practices Cyberseminar

Study Design and Data Decisions: Incorporating Genomics in Routine Care for Veterans with Colon Cancer

Sara J. Knight, PhD Birmingham VA Medical Center University of Alabama at Birmingham



February 14, 2017

The objectives of this session are:

To describe the use of qualitative and quantitative VA and non-VA data sources to characterize and investigate:

- The integration of genomic information into colorectal cancer (CRC) care in the VHA
- Barriers and facilitators to incorporating genomic information and services into routine care for Veterans diagnosed with CRC

Outline

- Background on genomic services for Veterans
- Previous research and resulting data decisions
- Retrospective cohort study: Methods, resulting data decisions and results
- Semi-structured interviews and content analysis
- Summary of lessons learned

Poll #3: Your experience with genomics

How many years of experience do you have working with genomic health services 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



Agenda

- Background on VA genomic services
- Preliminary research and resulting data decisions for this study
- Aim 1: Retrospective cohort study: Methods, resulting data decisions and results
- Aim 2: Semi-structured interviews and content analysis
- Summary of lessons learned

Agenda

- Background on VA genomic services
- Preliminary research and resulting data decisions for this study
- Aim 1: Retrospective cohort study: Methods, resulting data decisions and results
- Aim 2: Semi-structured interviews and content
 analysis
- Summary of lessons learned

Background: Genetic screening for colorectal cancer

- Lynch syndrome is uncommon (3% to 5% of all CRC)
- CRC is a common cancer
- Screening for Lynch syndrome among the first of all genetics applications to have accumulated substantial evidence for its clinical validity and utility
- Persons with the associated genetic mutation have a 80% lifetime risk of CRC and various levels of increased risk for seven to ten other cancers (e.g., small intestine, endometrium, ovary, stomach)

Background: Genetic screening for colorectal cancer (cont.)

- Longstanding guidelines recommended that all persons younger than age 50 diagnosed with CRC receive one or more genomic services (e.g., family history assessment, genetic counseling or consultation, molecular analysis, genetic testing)
- Limited knowledge of the patterns of care that have incorporated genomic information in VHA colon cancer screening and treatment

Genomic Services – Documentation Rates



Family Medical History



Genetic Testing



Genetic Counseling



Molecular Analysis

Lynch syndrome screening services



Figure 2. Adaptation of an algorithm for use in diagnosis of HNPCC.

- Family history documentation
- Genetic counseling or consultation with a clinical geneticist
- Analyze tumor for microsatellite instability (MSI) and/or
- immunohistochemistry (IHC)
- Genetic testing

Agenda

- Background on VA genomic services
- Preliminary research and resulting data decisions for this study
- Aim 1: Retrospective cohort study: Methods, resulting data decisions and results
- Aim 2: Semi-structured interviews and content
 analysis
- Summary of lessons learned

Preliminary studies

- Examined national VA administrative data 2003 2007 to describe documentation of genomic services delivered to Veterans under age 50 diagnosed with CRC and cared for in the VHA
 - Data sources: 2003 2007 VA administrative data, specifically VA outpatient, VA inpatient, and fee basis files
 - Variables: family history, genetic counseling, MSI/IHC, genetic testing measured using ICD-9 and CPT codes

Preliminary studies (cont.)

- Identification of Veterans age 50 and younger diagnosed with CRC (n=3,282) using CanCORS algorithm
- Documentation of family history, genetic counseling, or any molecular analysis and genetic testing was present for 6.7%, 0.12%, and 2.71%, respectively

Regional variation



Documentation of positive family history of colon cancer in Veterans diagnosed with colon cancer

Facility variation

- Positive family history of colon cancer among Veterans under age 50 diagnosed with colon cancer was low compared to the expected rate (20% of individuals with CRC have a positive family history of CRC based on prior epidemiological studies)
- No difference between VA medical centers affiliated with an academic medical center compared to those that had no academic affiliate (7% for each)

Facility variation (cont.)

- At the level of individual facilities, documented family history of CRC was consistently lower than what would be expected based on prior epidemiological research (approximately 20% with positive family history of CRC among individuals diagnosed with CRC)
- However, some facilities had documentation of family history higher than would be expected, suggesting both over reporting and under reporting
 - Among facilities reporting 10 or more cases of colon cancer, the highest rate of documentation was 40%
 - Among facilities reporting 20 or more cases of colon cancer, the highest rate of documentation was 22% (though the second-highest documenting facility was at 14%)

Data decisions based on preliminary work

- Aim 1
 - Necessary to use administrative data, clinical registry data, and the electronic medical record to fully characterize patterns of health services utilization
 - Importance of full text data from charts to use Natural Language processing to identify nascent services

Data decisions based on preliminary work

- Aim 1 (cont.)
 - Necessary to use administrative data, clinical registry data, and the AMA Physician Masterfile data needed to obtain complete data on physician characteristics (early key informant interviews suggested that physicians are primary adopters of Lynch syndrome screening)
 - Fee Basis files would be needed to document care outside the VA
 - New chart abstraction tool would need to be developed building from a tool developed for use in other health plans

Data decisions based on preliminary work (cont.)

- Aim 2
 - Facility variation of utilization of genomic services documentation could be used to stratify facilities for the barriers and facilitators study; Plan to select 15 high and 15 low genomic services utilizing facilities for key informant interviews

Conceptual framework

Hybrid conceptual model that integrated the Andersen behavioral model of access and the Greenhalgh diffusion of innovation model to inform adoption

- Innovation characteristics
- Adopter characteristics (including predisposing and enabling factors)
- Communication and influence
- System antecedents
- System readiness



Figure 1. Diffusion of Innovation Model for Health Services.

How is genomic information is incorporated into CRC care in the Veterans Health Administration?

The research question

Agenda

- Background on VA genomic services
- Preliminary research and resulting data decisions for this study
- Aim 1: Retrospective cohort study: Methods, resulting data decisions and results
- Aim 2: Semi-structured interviews and content
 analysis
- Summary of lessons learned

Mixed methods project using large VA and non-VA datasets and primary data collection

Aim 1: Retrospective cohort

How is genomic data incorporated into routine care of CRC patients

Aim 2: Interviews & content analysis

Barriers and facilitators of incorporating genomic info into routine care for CRC patients

Methods: Retrospective Cohort Design

- Used VA clinical and administrative data between 2003 and 2009
- Established cohort (n-866) using (ICD-9 and CPT codes) found in VA cancer registry and administrative data
- Adapted a chart abstraction tool for use in the VA based on previous study of utilization of genomic services in private health plans
- Four trained abstracters at the San Francisco VA Medical Center conducted chart review to collect data
- Obtained organizational characteristics of VA facilities from the 2009 National VA Oncology Facility Survey

Data Sources #1

Data sources: VA and non-VA	Source Office	Variables
National Patient Care Database Outpatient files	National Data Systems	Dependent variables, comorbidities, demographic information
National Patient Care Database Inpatient files	National Data Systems	Dependent variables, comorbidities, demographic information
Fee Basis Files	National Data Systems	Services from non-VA providers reimbursed by VA
VA Vital Status File	National Data Systems	Mortality – date of death

Data Sources #2

Data sources: VA and non-VA	Source Office	Variables
ADUSH Enrollment File	ADUSH-Office for Policy and Planning	Demographic information, patient eligibility for VA services, dates of enrollment, previous utilization of VA services
Planning System Support Group "Zip Code" file	PSSG	Patient travel distances to VA facilities
VHA Site Tracking System	PSSG or VSSC	Type of VA facility (full service, outpatient clinic)

Data Sources #3

Data sources: VA and non-VA	Source Office	Variables
Decision Support System National Data Extracts	Managerial Cost Accounting	Prescriptions, laboratory tests, test results
Oncology Facility Survey	Oncology Program Office	Oncology services by facility
VA Central Cancer Registry	Oncology Program Office	Clinical characteristics
External Peer Review Program (EPRP)	VHA Office of Quality and Performance	Clinical characteristics
American Medical Association (AMA) Physician Masterfile	Non-VA	Provider characteristics

Additional Data Source: VA National Electronic Health Record (EHR)

- National electronic medical record
- Two EHR data portals/platforms
 - Compensation & Pension Data Interchange (CAPRI)
 - VistAWeb
- Include full text data, e.g., notes
 - Dense and difficult to read
 - Extensive search capabilities available

VA Electronic Health Record (cont.)

- VistAWEB
 - Patient data across different facilities chronologically
 - Some text searching available
- CAPRI
 - Data from one single location shown at a time (not chronological)
 - Option to link to VistAWEB
 - Enhances search capabilities

Data decisions early in project

- CANCORS algorithm problems in this study—case finding problems and false positives
 - Algorithm developed for age 60 and older, did not work well in patients under age 50 most of whom were not diagnosed using CRC screening
 - Triangulating between the administrative data and cancer registry data critical in establishing the cohort
 - Change in plan of random sampling charts for abstraction because number of cases was fewer than 1000; Decision made to abstract all charts

Demographic and clinical characteristics of sample population

	N=886 (range or %)
Age	46 (19-55)
Male	812 (92)
Race	
White	518 (62)
Black	298 (35)
Tumor Stage	
0 to 1	196 (22)
2	196 (22)
3	227 (26)
4	181 (20)
Surgical resection	780 (88)
Chemotherapy	555 (63)

Rates of genomic services in Veterans diagnosed with CRC

Service	N=886 (%) ¹ 95% confidence interval %
Family history of colorectal cancer documented	188 (21.24) 18.59 - 24.09
Genetic counseling or consultation	43 (4.85) 3.53 – 6.48
Microsatellite Instability (MSI) or Immunohistochemistry (IHC) testing documented	71 (8.01) 6.31 - 10
Genetic testing for Lynch syndrome	12 (1.35) 0.70 - 2.35

¹ n=885 for family history, 1 case of missing data

Referrals for genetic counseling or consultation

Referral method	N=43 (%)
Clinician ordering referral for genetic counseling or consultation	
Oncologist	15 (35)
Gastroenterologist	12 (28)
Other health professional	5 (12)
Unknown	11 (26)
Genetic counseling or consultation completed and documented	28 (65)
Genetic counseling or consultation conducted in VA	8 (29)

Genomic medicine in facilities (n=108) with CRC cases (n=886)

	Genomic Medicine Event Documentation			
Overall GMEs	Family History	Counseling/ Consultation	Molecular Analysis	Genetic Testing
Facilities with zero GMEs N (%)	27 (25)	74 (69)	67 (62)	99 (92)
Facilities with GMEs Median (min – max)	2 (1 – 8)	1 (1 – 3)	1 (1 – 5)	1 (1 – 3)

Facility variation genetic counseling and consultation



Trend in referral for genetic consultation from 2003 through 2009



Trend in molecular analysis of tumor tissue from 2003 through 2009



Results to date

- How are genomic medicine services incorporated in CRC care in the VA?
 - Utilization of genomic services to identify Lynch syndrome in Veterans diagnosed with CRC was lower than expected based on established guidelines
 - Utilization of genomic services for CRC varied across VA facilities nationally from 2003 through 2009
 - Uptake in genomic services increased from 2003 through 2009
- Understanding organizational factors and individual patient and clinician perspectives on genomics services for CRC will be key to efforts to promote adoption

HSR&D Genomics Collaborator Events



VA Genomic Medicine Development

Decisions based on early analyses and change in VA policy during course of project

- VA Genomic Medicine Service was launched in 2012 approximately two years after the project was started
 - Genetic counseling using telehealth approach
 - MSI/IHC testing of tumor tissue of all Veterans diagnosed with CRC at any age and cared for in the VHA
- Expansion of cohort to include CRC patients of all ages and extending the time frame to 2014
 - Funding sought for this effort and was awarded

Agenda

- Background on VA genomic services
- Preliminary research and resulting data decisions for this study
- Aim 1: Retrospective cohort study: Methods, resulting data decisions and results
- Aim 2: Semi-structured interviews and content analysis
- Summary of lessons learned

Semi-structured interviews and content analysis (Key informant interviews)

 Aim 2: Identify barriers to and facilitators of the incorporation of genomic information and services into routine care for Veterans diagnosed with CRC

Barriers and facilitators





Article

Barriers and Facilitators to Adoption of Genomic Services for Colorectal Care within the Veterans Health Administration

Nina R. Sperber ^{1,2,*}, Sara M. Andrews ¹, Corrine I. Voils ^{1,2}, Gregory L. Green ³, Dawn Provenzale ⁴ and Sara Knight ^{5,6}

Key informant interview study

- Identify barriers to and facilitators of the incorporation of genomic information and services into routine care for Veterans diagnosed with CRC
- Selection of informants from high (n=15) and low (n=15) oncology volume facilities
- Conduct 100 to 300 semi-structured interviews by telephone with health professionals identified as knowledgeable about genomic services
- Expected interview time was 10 to 20 minutes
- Contact with Chief of Staff from each selected facility asked to identify clinicians and administrative staff members at their facility who were knowledgeable about colorectal cancer, Lynch syndrome, or genetic or genomic services

Table 2. Domains and constructs based on the Anderson Behavioral Model of Health Services¹ and the Diffusion of Innovation Model² used to analyze qualitative data on genomic services for colorectal cancer care in the Veterans Health Administration.

Domain	Construct	Definition
Structural	Availability ¹	Whether genomic services are perceived as being available at facility, regardless of whether or not used in-house
	Innovation-system fit ²	Fit with the organization's existing values, norms, goals, skill mix, ways of working; an aspect of system readiness for use of genomic services
	Incentives and mandates ²	Structural-level diagnostic and treatment guidelines, policies and procedures related to patient care
	Interorganizational networks ²	Linkages through common structures and explicit shared values and goals
Individual	Psychosocial factors (attitudes, knowledge) ³	Extent to which clinicians value incorporating genomics into colorectal care or demonstrate knowledge and interest; predisposing factor
	Enabling factors ^{1,3}	Resources that support clinicians' use of genomic services
Innovational	Relative advantage ²	Clear, unambiguous advantage in effectiveness of genomic services
	Augmentation/Support ²	Whether or not the genomic service comes with features to facilitate use, including templates, training, experts

¹ Phillips (1998, [15]); ² Greenhalgh (2004, [14]); ³ Bradley (2002, [13]); definitions adapted for this study.

Data decisions in key informant interview study

- Change in stratification criteria for facilities in key informant interview study
- Few facilities were found to have high utilization of genomic health services
- For this reason, we were unable to use high versus low genomic services utilization to stratify our sample for the barriers and facilitators interviews
- Decision to use high and low oncology volume to stratify sampling

Data decisions in key informant interview study (cont.)

- Decision that 58 interviews would provide sufficient information
- Sampling for maximum variation valuable in achieving comprehensive perspectives
- Interviews were longer than anticipated and informants were interested in the topic and questions
- Transcripts of the interviewers were much more informative then expected and saturation was reached as the sample size approached 50

Informant sample (n=58)

Specialty	n
Primary Care	11
Oncology	13
Surgery	7
Gastroenterology	18
Pathology	9

Barriers and facilitator findings

Key themes	Specific barriers/facilitators
Family history documentation routinely collected, but without guideline-informed policy or template and variable by specialty	 Lack of standardized template Too busy to add detail Interest in family history tools
Molecular tumor testing regarded as available and advantageous prior to genetic germline testing, though used seldom	 Clinicians see few cases of younger patients with CRC Lack of request and approval routines for referral Low cost of testing is advantage
Desire for expert support, such as genetic counseling, but variability in access to and awareness of this expertise	 Problems with approvals and lack of standard referral process Clinicians lack expertise to communicate with patients about genetics and Lynch specifically Perceive genetic experts as valuable

How were key informant interviews useful?

- Semi-structured interviews were conducted as a step toward planning implementation of genomic services for CRC
- Interviews provide an organizational context for understanding low documentation of genomic services, e.g., laboratory financial constraints limiting tumor tissue testing
- Identify targets for intervention, such as Electronic Health Record, clinician interest, lengthy approval processes, clinician lack of expertise

Agenda

- Background on VA genomic services
- Preliminary research and resulting data decisions for this study
- Aim 1: Retrospective cohort study: Methods, resulting data decisions and results
- Aim 2: Semi-structured interviews and content
 analysis
- Summary of lessons learned

Lessons Learned

- An earlier algorithm may not work in subsequent studies
 - Changes in guidelines, patterns of care, or disease and patient characteristics reduce the validity of algorithm inputs
- Best Practice: Early algorithm evaluation before and during a project is critical
- Data triangulation across multiple sources may reduce administrative data limitations
 - And can result in complex and unwieldy analytic data sets
- Best Practice: Early development of a priori plan for data integration can reduce the complexity and accelerate analyses and dissemination

Lessons Learned (cont.)

- Assumptions you made based on your preliminary studies, may not hold up in your larger project
 - Research funding allows for assembly of more complete data
 - Critical to be prepared to revise methods based on full data
- Best Practice: Team discussions as problems emerge and notes on the rationale for decisions are key!
- Diverse perspectives of research collaborators--an incredible gift in interdisciplinary science
 - Uncritical adoption of suggestions can limit the research or lead down an unproductive path
- Best Practice: Encourage vigorous discussion of issues, and then integrate discussion with prior evidence in making the final decision

Additional Resources

VA Data Resources



VIReC options for specific questions

HSRData Listserv

- Community knowledge sharing
- ~1,200 VA data users
- Researchers, operations, data stewards, managers
- Subscribe by visiting

http://vaww.virec.research.va.gov/Support/H SRData-L.htm (VA Intranet)



HelpDesk

Individualized support



Quick Guide: Resources for Using VA Data http://vaww.virec.research.va.gov/Toolkit/QG-Resources-for-Using-VA-Data.pdf (VA Intranet)

VIReC: <u>http://vaww.virec.research.va.gov/Index.htm</u> (VA Intranet)

VIReC Cyberseminars: <u>http://www.virec.research.va.gov/Resources/Cyberseminars.asp</u>

VHA Data Portal: <u>http://vaww.vhadataportal.med.va.gov/Home.aspx</u> (VA Intranet)

VINCI: <u>http://vaww.vinci.med.va.gov/vincicentral/</u> (VA Intranet)

Health Economics Resource Center (HERC): <u>http://vaww.herc.research.va.gov</u> (VA Intranet)

CDW: <u>https://vaww.cdw.va.gov/Pages/CDWHome.aspx</u> (VA Intranet)

Archived cyberseminar: What can the HSR&D Resource Centers do for you? http://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=101

Contact information

Sara Knight, PhD

Birmingham VA Medical Center

University of Alabama at Birmingham

sara.knight@va.gov

or sjknight@uab.edu



Next session: February 16, 2017

Good Data Practices Cyberseminar Series

Data Use and Data Decisions in a Mixed Methods Study about Hand Hygiene

Heather Reisinger, PhD

Center for Comprehensive Access & Delivery Research and Evaluation (CADRE), Iowa City VA Health Care System & University of Iowa Carver College of Medicine

