

Good Data Practices

Session 4

Research Application

Laurel A Copeland, PhD
September 12, 2013



Good Data Practices

- Series Recap
 - Session 1: Early Data Planning for Research
 - Session 2: Managing and Documenting Data Workflow
 - Session 3: Planning for Data Re-use

Session 4: Research Application

- Proposal planning and development
- Funding, IRB, and Study Initiation considerations
- Documentation Content and Locations
- Study Design and Implementation

Poll Question #1

- What is your primary research role?
 - Data analyst/programmer or statistician
 - Research coordinator or assistant
 - Student, trainee, or fellow
 - Investigator with data skills
 - Investigator without data skills
 - Manager, policy-maker, or other non-research VA stakeholder

Session 4: Outline

- **Proposal planning and development**
- Funding, IRB, and Study Initiation considerations
- Documentation Content and Locations
- Study Design and Implementation

Proposal planning and development

- Birth of a Research Project
 - During proposal development, the Aims take 1-3 months to develop
 - Aims inform Methods and Methods inform Aims
 - Funder is identified
 - Budget is proposed



Session 4: Outline

- Proposal planning and development
- **Funding, IRB, and Study Initiation considerations**
- Documentation Content and Locations
- Study Design and Implementation

Funding, IRB, and Study Initiation considerations

- Birth of Surgical Treatment Outcomes for Patients with Psychiatric Disorders (STOPP)
 1. **Systematic review** of the topic
 2. **Knowledge gap** identification
 3. **IRB protocol**
 4. **Pilot study funded**

Perioperative Outcomes and Safety in the Seriously Mentally Ill Elderly (POSSE)

Funding, IRB, and Study Initiation considerations

- Decision:
 - Begin the work before the funding, vs
 - Get the funds to do the work
- Draft a protocol for IRB:
 - Forms the basis of a fundable proposal
 - Allows work to begin
 - Time constraints on unfunded projects

Poll Question #2

- Which best describes your research experience with VA administrative databases?
 - Have worked directly with VA administrative data myself and am adept at DART process
 - Have worked directly with VA administrative data myself
 - Have collaborated on their analysis but have not pulled or manipulated them
 - Have not used VA administrative data at all

Funding, IRB, and Study Initiation considerations

- A Beautiful Trap: Unfunded Work
 - An IRB-approved protocol can generate:
 - Preliminary data
 - Baseline papers
 - Downside
 - Upside



Funding, IRB, and Study Initiation considerations

- **Pilot Study - Perioperative Outcomes and Safety in the Seriously Mentally Ill Elderly (POSSE)**
 - Had drafted a protocol for the IRB (LAC, VAL)
 - Wanted some funding to work on this IRB-approved protocol

Funding, IRB, and Study Initiation considerations

- Pilot Study – POSSE
 - Identified:
 - VISN 17 New Investigator Award mechanism
 - New Investigator – study team expands
 - Other team members: what capacity do you need?



Funding, IRB, and Study Initiation considerations

- Pilot Study – POSSE
 - Funding was awarded to J E Zeber for 2 years (2007-2009)
 - Plan: identify “**major surgery**” in FY2005
 - Defined “**inpatient surgery**”
 - Worked with surgeons and long lists of CPT and ICD9A codes iteratively to define major surgery

Session 4: Outline

- Proposal planning and development
- Funding, IRB, and Study Initiation considerations
- **Documentation Content and Locations**
- Study Design and Implementation

Documentation Content and Locations

- Pilot Study – POSSE
 - Met regularly
 - Documented meetings immediately with meeting minutes emailed to all team members
 - Documented all CPT and ICD9A code lists carefully with footers to identify what is listed and page numbers
 - Used good file names – concise but descriptive
 - Saved all documents to the Project Folder
 - Kept focused on the goal (define “**major surgery**”)

Documentation Content and Locations

- Project Folders & Binders
 - Research Oversight requires the PI to maintain a hard-copy Study Binder

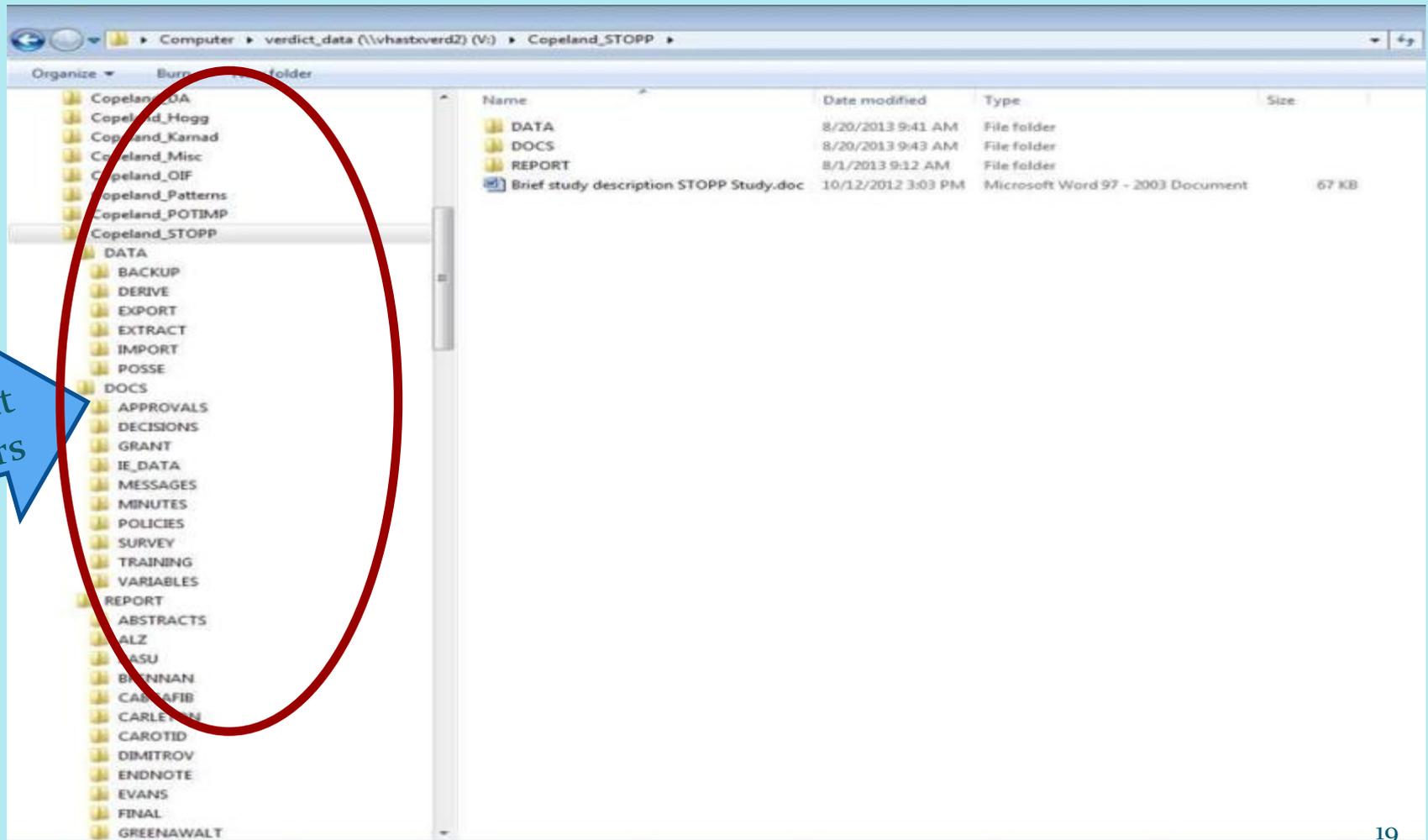


- Data Security requires VA Research Data to be stored within VA / in locations approved by IRB
- Reason demands an organized research project directory

Poll Question #3

- What level of experience do you have in organizing data, data structures, filenames, and variable names?
 - Expert, lots of experience, have worked with 1 or more schemas
 - Moderate, have used a schema for at least one study
 - Low, aware of the general idea but haven't implemented it
 - Novice, this is new to me

Documentation Content and Locations



Project folders

Documentation Content and Locations

- Project Folders
 - Naming conventions
 - Choose one model and stick with it to make it easy for team members to find code or output or paper drafts

\00412 Copeland STOPP

Or

\Copeland_STOPP

\sas_Copeland_STOPP

Documentation Content and Locations

- Project Documentation
 - Preserves decisions during study conduct and after study closure
 - Documented decisions inform manuscripts and reports
 - Documented processes allow replication
 - Documentation assists in data security, data management, and oversight/review



Documentation Content and Locations

- Project Documentation: Data
 - Save program files (SAS code) in a logical place (e.g., **Copeland_STOPP\DATA\DERIVE**)
 - Document the (SAS) program files
 - Filename
 - Date initiated
 - Programmer(s)
 - Brief study description including PI name
 - **Within the program, tell what you are doing and why**
 - The next programmer may not have access to the original programmer
 - **Study must be replicable**

Documentation Content and Locations

```
*prediabetes.sas*;
*28feb2008: this setup*; *1 a copeland*; *parent setup: schizdiabva.sas*;
    * SECTION I ANALYSIS
    * SECTION II BUILD DATASET
* VA HSRD IIR 05-326-02: Patterns of Care Veterans with Schizophrenia and/or Diabetes;
* DX Schizophrenia: 295xx omit 295.5;
* DX Diabetes Mellitus: 250.xx 357.2x, 362.0x, 366.4i;
* DATA SOU: PM/XM/NM, SE, LAR, PHA, BIRLS, VITALMINI, ENONEPER, SHEP - fy02-fy05;

LIBNAME SD '\\vhastxsas\sas_Copeland_Patterns\DATA'; LIBNAME shep '\\vhastxsas\sas_Copeland_Patterns\DATA\SHEP';
OPTIONS nofmterr nocenter errors=4 mergenoby=warn;
OPTIONS msglevel=i FORMCHAR='|-----+' formdlim=' ' ps=68 ls=100 nocenter; options font="Lucida console" 8;

*****;
**** SECTION I ****;
**** ANALYSIS: PRE-DIABETES STUDY ****;
*****;

DATA predm (DROP=priority);
MERGE sd.predm (DROP=priocat)
    sd.prediab_lac(IN=ini)
    sd.szdm(KEEP=scrssn SELIM_PHYS02 ptsd NCLASSVA02 AOD02-AOD05 prio use_after_death any: hispanic races black)
    sd.labs_gluld;
BY scrssn;
IF ini;

    * study inclusion / exclusion criteria *;
IF rxdiab=1 THEN DELETE; * 39904 ;
IF dmdx=1 THEN DELETE; * 39825 ;
IF prio>. ; * 39731 if first (so, 94) ...else 39711 ;
IF use_after_death ^=1; * 39804 if first (so 21) ...else 39711 (so an additional 20 over prio=.);
IF anytotclo2=1; * lose 485 ;
    * recodes *;
prio23678=(prio IN(2,3,6,7,8)); prio4=(prio=4); prio5=(prio=5);
testo2=MAX(fg02,a1c02); testo3=MAX(fg03,a1c03); testo4=MAX(fg04,a1c04); testo5=MAX(fg05,a1c05);
RUN; *this version has the followup aic/glu vars* 26sep08*17jul09* 264 vars 39226 was 39711 obs *;
```

Session 4: Outline

- Proposal planning and development
- Funding, IRB, and Study Initiation considerations
- Documentation Content and Locations
- **Study Design and Implementation**

Study Design and Implementation

- From Pilot to Merit Award
 - **Aims** – the larger study built on lessons learned from Pilot
 - **Methods** – informed by experience
 - **Preliminary data** – available from POSSE for STOPP proposal
 - **Weaknesses** of Pilot methods/staffing could be **addressed** in larger proposal

Study Design and Implementation

- STOPP Aims
 - **AIM 1:** Compare **rates** of surgery among VA patients by pre-existing severe mental illness status (SMI) status over the period FY06-FY09
 - **AIM 2:** Assess 30-day, 90-day, and 1-year postoperative **mortality** by pre-existing SMI status (covariate-adjusted models)
 - **AIM 3:** Assess 30-day, 90-day, and 1-year postoperative **complications** (MI, ACS, DVT-PE, ICU admission, pneumonia, respiratory failure, sepsis, wound infection, readmission)

Study Design and Implementation

- STOPP methods
 - Inpatient and outpatient files (**PM/XM, SE**): dx variables to define prior-year **Severe Mental Illness** (schizophrenia, bipolar disorder, PTSD, MDD)
 - Inpatient and outpatient files (**PM/XM, SE**): dx variables to define prior-year **comorbidity** (Charlson; Selim)
 - Inpatient and outpatient files (**PM/XM, SE, PS/XS, PP/XP**): procedure codes defined **Major Surgery** type(s), admission defined date of 1st qualifying op
 - Inpatient files (**PM/XM, PB/XB**): postoperative events were identified by dx variables (ICD9 codes) for 30-day, 90-day, and 1-year **postoperative complications** (MI, ACS, DVT-PE, ICU admission, pneumonia, respiratory failure, sepsis, wound infection, readmission)

Study Design and Implementation

- STOPP Methods
 - September **PSSG** files: contribute VA Priority Status [e.g., MDPPRD.MDP.SAS.NED.MAIN.SEP08.VSA.PSSG]
 - **MINI-VITALS**: Best Sex, Best DOB, Best DOB, and “use_after_dth_flg”
 - Inpatient and outpatient (**PM/XM, SF**) files: Period of Service
 - **OEFOIF ROSTER** file: OEF/OIF status (special DART application required)

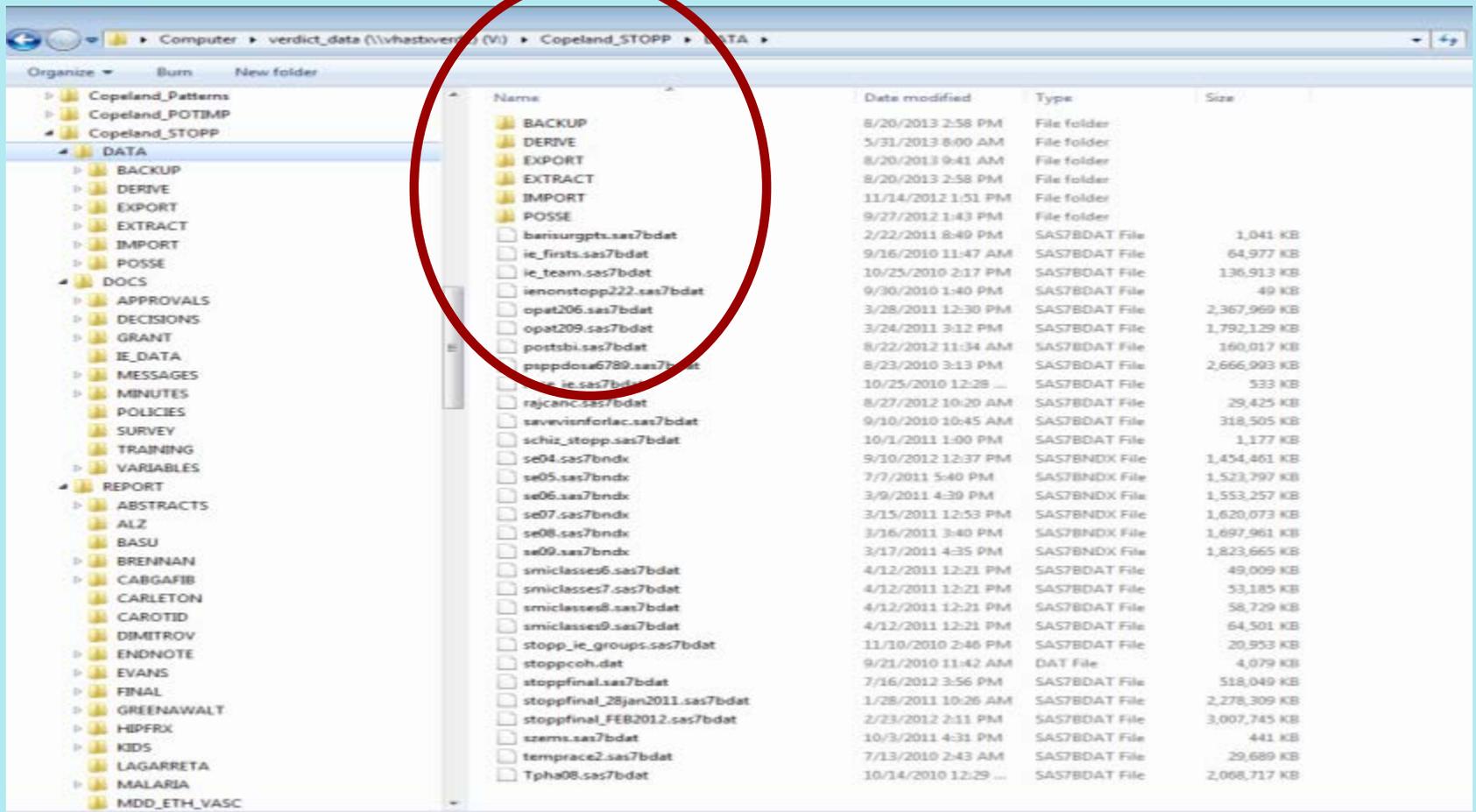
Study Design and Implementation

- STOPP Methods
 - DSS NDE **LAR** files: used to define lab-related process measures such as “had high cholesterol”
 - DSS NDE **PHA** files: used to define pharmacy-related covariates such as “using corticosteroids within 21 days prior to admission for surgery”
 - **NB:** DSS NDE’s are now incorporated into Corporate Data Warehouse (CDW) and are no longer generated as separate SAS datasets

Study Design and Implementation

- STOPP Methods
 - Used Encounter (**IE**) files to determine whether they could provide CPT codes for qualifying operations (they could not be so used)

Study Design and Implementation



Study Design and Implementation

- STOPP Methods
 - SAS copies of the flat files
 - Stored the binary versions at our “EXTRACT” level
(COMPRESS=YES)

Study Design and Implementation

- STOPP Methods
 - Also because of the size of the files, we converted algorithms for race, ethnicity, Selim comorbidity scores, and Charlson comorbidity index to **macros**
 - We experimented with various approaches to data aggregation and **asked HSRData listserv for input**

Study Design and Implementation

- STOPP Progress
 - Each paper has a sub-folder within \REPORT
 - All abstracts are in a single sub-folder within \REPORT

\Copeland_STOPP

\DATA

\DOCS

\REPORT

\ABSTRACTS

\MORTCARD

\SUICIDE

- Most papers needed a tailored analytic database as well, saved within \DATA (upper level)

Study Design and Implementation

- STOPP Progress
 - Team meetings to develop definitions, study designs, and papers
 - Data core meetings to monitor implementation of team decisions and generate patient-level dataset(s)
 - Aims determine primary papers
 - Mentees/colleagues determine secondary papers

Study Design and Implementation

- Single-site to Dual-site
 - STOPP funded April 2010
 - PI & co-Inv moved to a different VAMC in August 2010
 - **Prior** to the move:
 - Decided whether data would stay put
 - Filed an application with the new VA IRB
 - Began discussions with new VA's AO regarding moving funds

Study Design and Implementation

- A Dual-site Study Meets ORO
 - Or, how to attract unwanted attention...
 - ORO visited CTVHCS 3x in 2012
 - CTVHCS converted from using Policies to using SOP's
 - ORO reviewed our SOP's and our studies and requested many changes
 - ORO noted that some IRB-approved studies “sounded like” **data repositories** were being proposed
- I thought STOPP would be a good source of a Data Repository



Study Design and Implementation

- STOPP Data Repository
 - STOPP concluded funding in March 2013
 - Drafted the **SOP** and then **STOPP Data Repository Protocol**
 - Will permit use of the massive dataset to address aims not related to Surgical and Post-operative experiences of Veterans with Severe Mental Illness
 - **VINCI as a data storage location** was added to STOPP Protocol by IRB amendment and is specified in the Data Repository Protocol

Contact Information

Laurel A Copeland, PhD MPH

Center for Applied Health Research

A research center sponsored by Central Texas Veterans Health
Care System jointly with Scott & White Healthcare

2102 Birdcreek Drive

Temple TX 76502

Laurel.Copeland@va.gov

254-215-9880

Questions

Session 4: Research Application

- Recap
 - Proposal planning and development
 - Funding, IRB, and Study Initiation considerations
 - Documentation Content and Locations
 - Study Design and Implementation

Session 5: Research Application

- Preview
 - Planning for documentation of study design & measurement
 - Data cleaning
 - Construction of cohort
 - Outcomes construction
 - Covariate construction
 - Linkage of primary (survey) data & VA secondary data
 - Summary: Value of documentation