

How to Obtain Access to VA Electronic Health Data for a Research Study

So ... you've been assigned to coordinate a research study at South Texas Veterans Health Care System (STVHCS) that could benefit from access to VA electronic health record (EHR) data. It's the largest EHR database in the nation. Sounds great, right? But what do you do now?

This guide is designed to help you with this task in an efficient manner. It consists of two parts: "Quick Start/Refresher" guides that lists the basic steps to fill out either a Prep-to-Research request or a full-blown IRB Research request, and a "Complete" guide which includes screen shots and step-by-step instructions in case you need additional guidance for a specific step from the Quick Start guide. The Complete guide is recommended for first time data requesters.

This guide assumes you are either a VA employee at Audie L. Murphy VAMC, or you currently have WOC status with us, and you already have a VA email address and access to the VA intranet. If any of these are not true, then please check with your supervisor or office administrator to get those taken care of before trying to use this guide.

Valuable links for the data access request process:

- Preparatory to Research request process:
<http://vaww.vhadataportal.med.va.gov/DataAccess/PreptoResearchRequestProcess.aspx>
- IRB approved Research request process:
<http://vaww.vhadataportal.med.va.gov/DataAccess/DARTRequestProcess.aspx>
- Operations or Quality Improvement project process:
<http://vaww.vhadataportal.med.va.gov/DataAccess/OperationsAccess.aspx>
- Electronic application to request data for research purposes:
https://dart.vha.med.va.gov/vinci_dart_client/dart9/dashboard.html
- The official VHA guide for requesting data access:
http://vaww.vinci.med.va.gov/vincicentral/documents/DART_User_Guide.pdf
- Short 4-minute video on how to navigate the electronic application for PTR:
<https://vaww.vinci.med.va.gov/VinciCentral/VU/Video/193>
- Short 5-minute video on how to navigate the electronic application for IRB:
<https://vaww.vinci.med.va.gov/VinciCentral/VU/Video/199>

A short explanation of VHA data and who makes it available.

- Data enters the VA EHR through VistA, which is the VA medical software that VHA providers, nurses, pharmacists, and technicians use to document information during both inpatient and outpatient encounters, including date/time of visit, location of visit, vital signs, procedures done, diagnoses, laboratory test results, and medical prescriptions filled.
- The Business Intelligence Service Line (BISL), part of the VA Office of Information and Technology (OI&T), is responsible for converting the data in VistA into a database format called the Corporate Data Warehouse (CDW) which can be queried by data analysts.
- The VA Informatics and Computing Infrastructure (VINCI) provides a platform for access to and analysis of VA data, which includes the CDW. VINCI has created an online application process called Data Access Review Tracker (DART), which is used to request data for research purposes.

- National Data Systems (NDS) is the organization that processes your DART request, primarily checking to see that the paperwork has been filled out correctly and making sure all the correct agencies do their part to verify that everyone listed on the DART request meets the data access standards.
- The VA Information Resource Center (VIREC) provides training and written products that guide VHA data users on what's available in the CDW and other data sources, and how to use it efficiently.

How to use this guide. Let's start with some quick questions to make sure this guide is for you, and to direct you to the parts of the guide you need.

1. Are you or is someone on your team trained in the SQL database language and/or in SAS data processing language?
 - a. If yes, then follow this guide to get you through the DART process, and then ask your analyst to complete the query.
 - b. If not, don't worry. There are data analysts in VERDICT who know SQL and/or SAS who can do most of the data querying for you. Ask the STVHCS Research Service to refer you to one of the analysts. You should meet the analyst prior to the DART request because he/she must be listed in the DART application, and may also be able to help with the application.
2. Is your study Research (as opposed to Operations)?
 - a. If your study is intended to contribute to generalizable knowledge or it is/will be funded as research, then it is a Research project.
 - b. If you already have IRB and R&D approval letters, then your study is Research.
 - c. Your study may qualify as Operations if your team will be using EHR data for VHA program administration, quality improvement, or program evaluation (basically, using data to evaluate how well an important health care process or activity is going). If you think the study qualifies as Operations, you should confirm this through the STVHCS Research Service.
3. Which research function are you gathering information for?
 - a. You are working on a research proposal to determine the feasibility of conducting the study (such as the number of patients that meet the inclusion and exclusion criteria):
Follow the Preparatory to Research process, Section C, starting on p. 3.
 - b. You have IRB and R&D Committee approvals and would like a patient list for recruitment, or you will do a retrospective study with EHR data:
Follow the IRB Research process, Section D, starting on p. 12.
4. For PTR queries, can you get it done without personally identifiable information (PII), such as patient name, SSN, date of birth, address?
 - a. If you don't need to search through CPRS to verify inclusion or exclusion criteria, save time and hassle by not requesting Real or Scrambled SSN. The data analyst can search records and count unique patients with the patient identifiers called PatientSID and PatientICN.
 - b. If you need PII, even if it's just last name and last 4 of SSN to aid with looking up records in CPRS, then the DART application must request Real SSN, which will require more paperwork and a few more days to process.

Explanatory note for the data analyst new to the CDW:

All EHR database queries need a patient identifier to be able to connect data from different tables, and to ensure there is just one set of records for each unique person. The VistA system assigns a unique identifier (**PatientIEN**) for each new patient within its 3-digit station, called Sta3n (such as 671 for South Texas or 674 for Central Texas). A patient at Audie L. Murphy VAMC who also has appointments at Frank M. Tejada Outpatient Clinic has the same PatientIEN in both places. However, PatientIEN is only unique within a specific Sta3n, so a certain value of PatientIEN in the STVHCS is probably used for a different patient in the Central Texas Health Care System in Temple and for yet again a different patient in the Texas Valley Coastal Bend system in Harlingen. At the other end of the spectrum, VHA patients are free to visit more than one VA health care system (Sta3n), so they will have different PatientIENs in each VistA system. To prevent confusion and assure uniqueness of patient identifiers within the CDW, two different types of numbers are assigned. One is the **PatientSID**, which is unique for each combination of patient and VistA station, and it is the primary key (unique identifier) used for many of the tables in the CDW database. No two persons have the same PatientSID anywhere in the VA national EHR system. However, if a person has visited more than one VA health care system (e.g., someone with VHA appointments in both San Antonio and Dallas) that person will have a different PatientSID for each different system (Sta3n). To make it easy to consolidate records for patients who have visited more than one system, the CDW makes available the **PatientICN**, which can be matched to all the PatientSIDs of the same person. Business logic is used to collect together all the PatientSIDs that share the same social security number, name, and date of birth, and assign them to one PatientICN. The CDW has a crosswalk table for matching PatientICN to the unique combination of PatientSID and Sta3n.

If You Want to Learn How to Do Some or All of This on Your Own.

1. Cyberseminar on Requesting Access to VA Data (11/5/2018). About 1 hour long.
https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=3540
2. Other VIREC cyberseminars, with a general description and a link to the archive:
<https://www.virec.research.va.gov/Resources/Cyberseminars.asp>
3. Join the Health Services Research & Development (HSR&D) Listserv, which encourages VA data users to ask and answer questions about VA data:
<https://vaww.virec.research.va.gov/Support/HSRData-L.htm>
4. Learn about the types of data available, and how to get started writing SQL code:
<http://vaww.virec.research.va.gov/CDW/Documentation.htm>

(For data analysts:

CDW guide Introduction and Policies <https://vaww.cdw.va.gov/Pages/CDWHome.aspx> ,

MetaData report <https://vaww.cdw.va.gov/metadata/default.aspx>

VHA Data Portal <http://vaww.vhadatportal.med.va.gov/Home.aspx>)