

Women's Health Initiative

Accessing and Analyzing WHI Data Using the Virtual Data Enclave (VDE)

Overview

The Women's Health Initiative Clinical Coordinating Center (WHI CCC) developed a Virtual Data Enclave (VDE) to expand access to selected confidential data for those who meet approved criteria. The VDE allows Investigators with an approved manuscript or ancillary study proposal to remotely access and analyze data on a secure server located at the CCC.

Access to the VDE will be granted to selected investigators for the purpose of increasing the knowledge gained from the WHI resource. Users of these data are expected to respect the confidentiality of participant information and the tremendous investment of NHLBI and the WHI investigators in developing this resource by using it in accordance with all WHI policies. Access to these data may be terminated by the CCC at any time if the user fails to abide by the data protections outlined in the Data Use Agreement.

Access to data in the VDE is based on the requirements of the specific manuscript or ancillary study proposal. Data in the VDE do not include direct identifiers (e.g., name, social security number). Investigators using the VDE are able to view but not print or save data, files, tables, or screen shots to their local computers. CCC staff will review and redact as necessary any files requested to be transferred outside of the VDE to ensure they meet WHI confidentiality requirements prior to releasing them to the investigator.

Data Available in the VDE

- **Participant medical history and medical services use.** Includes diagnoses, medication use, medical provider stays, and medical services usage, dates, and costs and location of services obtained from Medicare (CMS) data linkage, the National Death Index and any other data linkage activities that may be approved and conducted in the future
- **Participant residential history.** Includes dates of residence along with latitude, longitude, FIPS code, and zip code of residence since being entered into the WHI database. Actual street addresses will not be included. State/country of birth and state/country of residence at various ages may be included upon request.

Please note: Because of the confidentiality protections required for Medicare data, projects requiring linkage of address history or geospatial information with Medicare files are not allowed.

- **Participant's responsible Clinical Center history.** Includes Clinical Center ID and the dates the Clinical Center was responsible for the participant over time.
- **Participant vital statistics.** Includes date of birth and date of death.
- **Other participant study related information.** Includes dates of enrollment/randomization, dates of consent, and dates of participant contact for collection of data and specimens.

Data Excluded from the VDE

- Participant name, residential street addresses, phone numbers, and Social Security Number are not available in the VDE.

Cost

The standard VDE fee is \$3,000. The fee covers the creation, management and regulatory compliance of your account. Additional charges may be necessary for studies requiring additional review, services, or access. Considerations include the data to be accessed, the timeline, and the amount of CCC support needed.

Requesting Access to the VDE

- Prepare a manuscript or ancillary study proposal and obtain approval through the WHI Publications & Presentations (P&P) Committee (manuscript proposals) or WHI Ancillary Study (AS) Committee (study proposals). Details on the requirements and process for developing and submitting a proposal are available at www.whi.org. The specific data files or sensitive data elements needed should be specified in the proposal.
- Complete the *WHI Virtual Data Enclave (VDE) Intake Form*, identifying the parameters of the study or manuscript, specific data files needed, statistical software to be used, and project timeframe, and submit it to the CCC. Proposals that involve the use of CMS data are submitted for an additional review by the CCC to confirm availability of the requested data and assess whether additional statistical expertise may be needed before analyzing these complex data.
- Submit the following to the CCC:
 1. A signed *WHI CMS Data Use Agreement (VDE CMS DUA)* (for proposals using the VDE to analyze CMS data) or *WHI VDE Data Use Agreement (VDE DUA)* (for proposals using the VDE to analyze data other than CMS data). These documents address confidentiality and security issues related to the sensitive data available through the VDE. (If you intend to designate an analyst to use the VDE, they will also be required to sign the agreement.) Note that the CMS DUA requires institutional signatures, which may require additional time for review and approval at your institution, impacting your overall analysis timeline.
 2. Documentation of local IRB approval for these activities. The approval should note the confidential information that the study seeks to access.
 3. A Purchase Order (PO) from your institution in the amount specified in the previously provided quote.
- Upon receipt of the signed agreements and PO, the CCC will:
 1. Process the PO (funds must be provided in advance of use of the VDE).
 2. Process and obtain IRB approval from the FHCRC to provide VDE access to the investigator.
 3. Create a user ID and password authorizing access to the specific set of data files needed to achieve the scientific goals of the approved proposal.

Using the VDE

Please refer to *The WHI Virtual Data Enclave User Guide.docx* at www.whi.org.

Additional Considerations

- When the timeframe established in the funding agreement ends, access to the VDE will be terminated (typically one year). You must have obtained all desired data summaries and/or analysis results before the timeframe closes, or request an extension. The timeframe of your VDE access begins when the CCC provides the study-specific VDE account information to you.
- If you need more time or CCC assistance (e.g., programming or statistical consultation) than originally planned, please contact the CCC to determine if additional funding is needed to accommodate the change and if so, to obtain a new quote.
- If you need access to additional data, the CCC will review your request to determine if it falls within your originally approved aims. If not, you will need to obtain approval from the P&P or AS committee, your local IRB, and the FHCRC IRB before access to the additional data is granted.
- If the CCC has reason to believe the investigator may be misusing VDE access, the FHCRC Legal department will be immediately contacted for appropriate next steps.

Frequently Asked Questions

- *Who can have access to the VDE?*
Any investigator with an approved manuscript or ancillary study proposal, with local IRB approval, and who has signed the necessary agreements can access the VDE.
- *To what data can I get access?*
Data currently available on the VDE includes all WHI data currently available on the WHI website, Medicare data for years 1991-2016 (updated annually approximately 15 months after the end of the calendar year), and some confidential WHI data, such as geocoded addresses, other information on geographical location (e.g., zip codes), and calendar dates (e.g., enrollment dates). **Because of the confidentiality protections required for Medicare data, projects requiring linkage of address history or geospatial information with Medicare files are not allowed.**
- *I'm not sure which specific Medicare or WHI files I need – any suggested resources?*
When developing a proposal and/or completing the intake form, consult www.whi.org for details on WHI data documentation and www.resdac.org for CMS file descriptions, data documentation, variable lists, and SAS code. *Can I link additional data to the WHI data?*
It depends on the data. This requires additional discussion in advance with the CCC.
- *What statistical software is available in the VDE?*
SAS, STATA, and R are available in the VDE.
- *What if I need other software?*
Installing other software in the VDE is not supported.

- *How long will it take to get my account set up?*
Once approvals are in place, we typically will provide your account information within one week. Access will be granted for one year.
- *How will the access work?*
You will access the VDE using the Remote Desktop Connection feature in Microsoft Windows and the username, password and server location provided by the CCC.
- *Can I access the VDE using a Mac?*
Yes. There is a free download available for Macs that allows you to access the VDE using Remote Desktop Connection.
- *How long will it take for the CCC to review my “take away” files?*
The CCC will try to review and provide your output within one week, however this will depend in part on the volume of your files.
- *How will I get my files?*
Depending on the content and size of the files, they may be emailed or transferred via a secure file transfer site.
- *Can I create data files to take out of the VDE?*
While this is not currently supported, it is possible that a data file with all protected information removed could be released. Consult the CCC to discuss the feasibility and cost.
- *How much will it cost?*
The standard VDE access fee is \$3,000, which covers the creation, management and regulatory compliance of your account. Additional charges may be necessary for studies requiring additional review, services, or access.
- *What if I haven't finished my project at the end of the one year of access?*
Please contact the CCC with an extension request and justification (email Lindsey Bull at lbull@WHI.org).