Data Access Using the VDE
April 2020
Why are NAHDAP data in the ICPSR VDE?

The National Addiction & HIV Data Archive Program (NAHDAP), funded by the National Institute on Drug Abuse (NIDA), is hosted at the Inter-university Consortium for Political and Social Research (ICPSR). NAHDAP data in the ICPSR Virtual Data Enclave (VDE) are restricted from general dissemination to protect the confidentiality of the individuals and/or organizations represented in the data. To access data in the VDE, a Restricted Data Use Agreement (RDUA) must be established between the University of Michigan and the researcher's institution.

How does it work?

A link to the online application is provided in the Access Notes of the online study description for the restricted-use dataset. Researchers interested in accessing the restricted-use data using the VDE must complete the online application. After the researcher creates an online project, the RDUA is sent by email to the researcher. The researcher returns the signed RDUA to NAHDAP by email. NOTE: Researchers on the same project but at different institutions must each provide a signed RDUA – one from each institution – because this agreement is a legal document with the researcher’s institution, not the researcher him or herself.

Once a request is approved, researchers and their team members can download the required software client. Documentation will then be provided for them to become familiar with the VDE system.

When a researcher logs on to the VDE, a virtual machine is launched on the researcher's own desktop but operates from a server at ICPSR. The virtual machine is isolated from the researcher's physical desktop computer – users cannot download or upload files or parts of files from or to the VDE; print VDE contents to a printer; or email, copy, or otherwise move files in or out of the VDE computing environment, either accidentally or intentionally.

For the researcher, the VDE operates just like a standard Windows desktop. All work with the study files can only be done within the VDE. Data files and software remain on the ICPSR server. Statistical analysis packages available for use within the VDE include:

- SAS 9.3
- SPSS 19
- Stata 12
- StatTransfer 9
- SUDAAN (SAS-callable) 10
- R 2.15 (CRAN repository mirror)
- HLM
- MPlus
- WinBugs
- ArcGIS 10
- Adobe Acrobat
- Microsoft Office 2010
- Emacs 21
- Textpad 5
- WinZip 14
Making a request

1. **Researcher** selects *Apply for NAHDAP* under the *User Menu* and creates an online project with:
   - Information about the project and why the restricted-use data are needed
   - Investigator’s contact information
   - List of data files requested
2. **NAHDAP’s** System displays that a project is created and sends email with blank Restricted Data Use Agreement (RDUA)
3. **Researcher** sends an email with:
   - RDUA signed by the Investigator and Institutional Representative
   - Agreement to Data Security Plan procedures (as part of the RDUA)
   - IRB documentation (approval or exemption)
   - CV/résumé/NIH biosketch for each user that will access the VDE
4. **NAHDAP staff** upload files, review the request, and may return the request to Researcher if more information is required
5. **Researcher** makes the requested changes, resubmits the request, and pays the access fee if a fee is required to access that dataset
6. **NAHDAP staff** approves the request
7. **Researcher** assigns user licenses for VDE access
8. **ICPSR computer staff** configure University of Michigan accounts for each research team members granted access to the data and sends instructions on how to access the VDE

Making a request: Frequently Asked Questions

1. **Who is eligible to access the data?**

   According to the RDUA, the Investigator is the researcher primarily responsible for analysis and other use of the files in the VDE; assumes the responsibility of completing the online access request and providing required documents, reports, and amendments; and agrees to responsibly manage and use the data in the VDE (including the work of all research team members) and implement all security procedures per the Data Security Plan.

   The Investigator must demonstrate experience with using restricted-use data (defined as Confidential Data in the RDUA) by providing information that summarizes their recent/relevant research, including references. A student or nonstudent Investigator with no prior experience using restricted-use data should reference the research of the Sponsor that he or she has.
participated in if this activity is germane to show an understanding of how to use restricted-use data or provide recent/relevant research of the Sponsor who has demonstrated experience. This information is entered in the Project Description box in the online request system.

2. **Does the Investigator need to complete the request herself or himself?**

   The Investigator can designate someone to enter the information into the online system and provide the needed documents. The automated emails from ICPSR regarding the request will be sent to the person named as Project Administrator on the Create Project screen.

3. **What needs to be provided in the Project Description box of the online request system?**

   The description is limited to 3800 characters. Most important is describing the project as research with a short reference to theory or historical context, how the project extends what is presently known, and why the restricted-use data are needed (rather than other data) to do the project. Each of the study’s datasets (or reasonable group of datasets) needed should be listed along with information on why and how the dataset will be used.

4. **I clicked on the Create Project button in the online request system but need to add more information. What should I do?**

   Requests are submitted when the Create Button is used. Please send email to nahdap-vde-support@umich.edu to have your project returned to you for you to edit your information. Please do not start another Project.

5. **What is the cost of data access?**

   For the Population Assessment of Tobacco and Health (PATH) Study restricted-use data, the National Institute on Drug Abuse (NIDA) is providing a pool of virtual machines. Users of the VDE do not pay a fee to access these data.

   For other NAHDAP restricted-use data available in the VDE, project teams are charged an annual fee of $405/user.

   Each research team member with access to the data is required to have an account.

6. **How long does the request process take?**

   NAHDAP staff typically review initial requests within 1-3 business days after submission. The length of the review process depends on whether it is necessary to return the request to the investigator for modification or more information. Requested modifications to the RDUA must be approved by the University of Michigan and will extend the review timeframe (see Making a Request: FAQ #11). Following approval of the completed request, the Investigator can expect access within 1-2 weeks.
7. **How long is my period of data access?**

The access period is two years. Researchers needing more time to complete their project can request an extension for another year. If doing so, they also need to make sure all information submitted previously is still current.

If a member of the research team changes institutions, a new RDUA must be established with his or her new institution before the individual can continue to access the data in the VDE. A change in institution by the Investigator terminates the RDUA for all VDE users. If remaining VDE users want to continue access, a new Investigator needs to be designated and a new RDUA submitted signed by the new Investigator and institution. (See Using the VDE: FAQs #3 for details).

8. **The RDUA must be signed by the Investigator and an Institutional Representative. Who is my Institutional Representative?**

Your Institutional Representative is an individual who has the authority to enter your institution into legally binding agreements. The Restricted Data Use Agreement is approved by the University of Michigan Office of Research and Sponsored Projects. Although an institution-to-institution agreement, the RDUA is unilateral and only signed by the Investigator and a legal representative of his or her Institution.

9. **Can individuals at different institutions collaborate on one research project?**

Yes. The ability for researchers at different institutions to collaborate is one of the main advantages of using the VDE. Please note, however, that RDUA is between the University of Michigan and the Investigator's institution. Therefore, if two researchers are on the project and are affiliated with different institutions, two RDUA must be executed – one for each institution.

10. **My institution does not have an Institutional Review Board (IRB). Do I still need to provide an IRB document?**

Yes. Institutions without an IRB should have the project reviewed by an accredited organization listed with the Association for the Accreditation of Human Research Protection Program, Inc. (AAHRPP). Their website is [http://www.aahrpp.org/learn/find-an-accredited-organization](http://www.aahrpp.org/learn/find-an-accredited-organization). Investigators should submit to NAHDAP the approval or exemption document received from the accredited organization. Please send questions or problems with this process to nahdap-vde-support@umich.edu.

11. **My institution would like to propose modifications to the Restricted Data Use Agreement. Is this possible?**

The terms and conditions of the RDUA are written to apply uniformly to all data users and without modification. An Institution can request modifications to the RDUA by preparing a version of the RDUA in which the requested changes are tracked or otherwise clearly identified. NAHDAP staff will send RDUA with requested modifications to the University of
Michigan Office of Research and Sponsored Projects (UM ORSP) for approval, negotiation, or disapproval. This process will extend the time to obtain approval to access the VDE. The UM ORSP is under no obligation to approve or negotiate requested changes to the RDUA and retains the right to disapprove any such changes.

12. Can I access the VDE from my home office?

Working on a laptop or working from home is possible, as long as work is completed in an office that meets the requirements of a Secure Project Office, which are listed below:

- The computer or monitor screen is not visible from open doors or through windows
- The office door must be closed when you are logged into the VDE
- Only approved research team members can be in the office when you are logged into the VDE
- While logged into the VDE, if you and any other research team members are out of the office, the door must be locked

13. Can I request that additional software be added?

The VDE offers a wide range of software for researchers to use (see page 2). However, requests for additional software to be installed in the VDE are reviewed every 3 months. Request should be emailed to nahdap@icpsr.umich.edu. Please note that requested software packages with little demand are unlikely to be added.

14. Can I import files into the VDE?

Researchers interested in having their own files (e.g., existing syntax) for use in the VDE with the requested restricted-use data can email the files to nahdap@icpsr.umich.edu. ICPSR computer staff will vet the files for any unallowable content before researcher-supplied files are imported into the VDE.

15. What is the Duo Two-Factor Account and for what devices can it be used?

To log into the VDE requires a University of Michigan uniqname (login) and password and a second-factor device from Duo Security. Duo provides a smartphone app for iOS, Android, Blackberry, and Windows phones. ICPSR will send instructions for the Duo app once a user’s VDE account is ready. A physical token in the form of a key fob can be purchased from ICPSR. Please factor into your project timeline for a physical token to be sent to you by Certified Mail.

16. Why do I use an ICPSR MyData account to request the data but use a UM account to access the VDE?

At this time, the ICPSR VDE Management System requires use of an ICPSR MyData account. The VDE uses the University of Michigan Active Directory because the VDE resides on the University of Michigan system. ICPSR cannot use Active Directory for the VDE Management System because potential VDE users do not have a University of Michigan account at the time of application. ICPSR MyData cannot be used for the VDE because it does not have access to the University of Michigan computing system.
Using the VDE

ICPSR computer staff configure University of Michigan accounts for each researcher on the project named as a VDE user in the online system. The researcher obtains the smartphone app from Duo Security for second-factor authentication needed for login (see Making a Request: FAQ #15 for details).

The researcher also downloads the required VMWare software and uses the provided documentation about the VDE access process and environment to log into the VDE. All researchers electronically sign a confidentiality pledge each time they access the VDE.

Researchers use the VDE within their project space as they would with their own desktop computer, e.g., create folders, run statistical programs, and edit documents.

All output for review is placed in a “DisclosureReview” folder and prepared according to the instructions in the vetting requirements document in this folder. Output can include Word tables, Excel spreadsheets, and graphics (preferred) or raw output from a statistical package (files and logs). Requests for output review are sent to nahdap-vde-support@umich.edu.

Using the VDE: Frequently Asked Questions

1. Can I record information outside the VDE?
   
   No. As a user of restricted-use data you are obligated to protect the confidentiality of individuals and organization represented in the data. No information may be recorded (e.g., no writing on paper or typing into editing software) outside of the VDE. Any information to have outside the VDE should be electronically stored in files in the “DisclosureReview” folder in the project’s space in the VDE. NAHDAP staff will conduct a disclosure review of the files and provided them to the user if cleared of disclosure concerns.

2. I need to change my office location. Do I need to let you know about it?
   
   Changes in office locations are not required to be reported unless the new office is outside the purview of the Institution that is a party to the RDUA. Please note that the new office location must still meet the criteria for a Secure Project Office as defined in the Data Security Plan (see also Making a Request: FAQ #12).

3. I have a current Restricted Data Use Agreement, but I am changing institutions. Can I continue to work on this project at my new Institution?
   
   Yes, but in order to continue using the data in the VDE a new RDUA must be executed with the new institution first. Instructions are in Section VI. of the RDUA, specifically:

   Notify NAHDAP/ICPSR of any change in the institutional affiliation of the Investigator or any VDE User, or the addition or removal of a VDE User on the research project and need to access the restricted-use data. Notification must be in writing and must
be received by NAHDAP/ICPSR at least six (6) weeks prior to the last day of employment with Institution or as soon as possible for a change in membership on the research project. Investigator’s separation from Institution terminates this RDUA. Investigator may reapply for access to the restricted-use data as an employee of the new institution. Re-application requires:

- Execution of a new RDUA signed by both the Investigator and the proposed new institution;
- Information about any VDE Users at the proposed new institution;
- Agreement with the Data Security Plan (which is Attachment A of the RDUA); and
- IRB document with approval or exemption of the research project at the proposed new Institution.

These materials must be approved by NAHDAP/ICPSR before access to the restricted-use data and any derivatives or analyses stored in the ICPSR VDE can be provided to Investigator after Investigator’s change in Institution occurs.

4. **Are RDUAs transferable if a researcher leaves the team and another takes his/her place?**

   If a research team member no longer needs access to the restricted-use data in the VDE, the Investigator (or designee) should send email to nahdap-vde-support@umich.edu to request removal of the user from the VDE. The user’s license can be transferred to a new research team member by adding the new member’s information in the online system so an account can be configured. User licenses are active for one year and expire based on the date when originally assigned. Per the Data Security Plan, VDE users cannot share or give their VDE login and password to anyone, including with other VDE users on the same research project. If the Investigator leaves the Institution, the RDUA is terminated. Researchers remaining at the Institution who want to continue access to the VDE must execute a new RDUA.

5. **Can I have files (e.g., programming code) uploaded to the VDE?**

   Yes. Email your file to nahdap-vde-support@umich.edu. After required checks, NAHDAP staff will upload the file into your VDE workspace and notify you where it is located.

6. **How do I amend my approved project?**

   Investigators that want to extend the scope of their project or request additional waves or years for the data they are using should send their request to nahdap-vde-support@umich.edu.

7. **How do I add users to my project so they get a VDE account?**

   Investigators should use the link in the email that notified them of their approval to return to the VDE Management System and provide the new users information. Please also notify NAHDAP staff by email to nahdap-vde-support@umich.edu. New users must be affiliated with an Institution that provided a RDUA for the project. If not, a RDUA must also be submitted to nahdap-vde-support@umich.edu.
Obtaining VDE Output

1. **Researcher** ensures all output for review is in the “DisclosureReview” folder (see details in Obtaining VDE Output: FAQs below)

2. **Researcher** (or designee) submits a request for output review through Footprints (correspondence tracking software) by sending email to nahdap-vde-support@umich.edu

3. **NAHDAP** or ICPSR staff reviews the output

4. **NAHDAP staff** may send output that cleared disclosure review or request modifications prior to a second review. Correspondence is sent through Footprints

5. **Researcher** completes any requested changes and resubmits the request.

6. **NAHDAP staff** review the request, approve the output, and notify the Researcher through Footprints

7. **NAHDAP staff** send the output to the Researcher and anyone copied on the original Footprints request.

ICPSR archives project files for 3 years after the project ends.

Obtaining VDE Output: Frequently Asked Questions

1. **How will I remove output when I’m ready to present my findings?**

   Researchers interested in removing output or other non-data files from the VDE may submit a request to NAHDAP, and staff will review the requested files to ensure no confidential information is removed from the VDE. Requirements for a specific dataset will be in the “DisclosureReview” folder in the VDE. *Never* make the requirements public.

   Researchers should send the following information to nahdap-vde-support@umich.edu:

   - A narrative summary of all output requested for release, compiled into a single document, one-page maximum.
   - File names of all documents requested.
   - Weighted output should be accompanied by unweighted output to ensure both meet the vetting requirements.
2. **What disclosure review criteria are used to vet my output?**

General vetting requirements and criteria to protect the confidentiality of individuals and/or organizations are provided below. The disclosure review procedures and vetting requirements for a specific dataset will be available to the research project team in their project space within the VDE. *Never* make the vetting requirements public.

Note: all output for review must be placed in the “DisclosureReview” folder in the project space in the VDE. Reviews are for disclosure risk as well as for malware. If the approved output includes multiple files, a zip package containing all the files is made and provided.

- Well-annotated program files need to accompany all output.
- No data files, data extracts, or listing of cases will be released.
- No output broken out into subsets less than the number of cases specified as the minimum for the specific dataset will be allowed.
- **Tabular output:**
  - Tables must include cell counts and percentages.
  - Tables with cell sizes smaller than the threshold for the specific dataset will not be released. The researcher is responsible for recoding variables as necessary before submitting tables for review.
  - Tables must be accompanied by a relevant subset count and description.
  - Multiple or linked tables from the data or subsample will be reviewed for “disclosure by differencing” in which subtraction or other type of calculations can produce results below the required threshold.
- **Model output:**
  - Model output must include a description of any sub sampling of variables included in the model.
  - Model output based on fewer than the number of cases specified as the minimum for the specific dataset will not be released.
- **Weighted output:**
  - If output for release must be weighted as stated in the documentation, both unweighted and weighted output should be provided for vetting to ensure the unweighted counts met the required thresholds.
  - If output for release is not required to be weighted as stated in the documentation, the unweighted output must meet the disclosure review criteria above.

Requests for output are limited to results needed for presentations and papers. Reviews of requests for output will vary in length and in proportion to the size and scope of the requested output.

3. **How long should I expect the vetting of my output should take?**

The target review period is 10 business days for each output review request. Keep in mind that reviews can involve iterations of review and output changes before output is cleared by the
disclosure review process and able to be sent to the researcher. The researcher’s project timeline should allow for some iteration of reviews.

Each request is reviewed in the order it is received. For this reason, users are discouraged from submitting multiple requests at one time. Multiple requests will result in delays, as NAHDAP will review them in order and one at a time.

4. **What can I do with the output that is sent to me after it is vetted?**

Output within the VDE can only be shared with other VDE users sharing the project VDE workspace. Output sent to a researcher after being vetted for disclosure risk is considered public-use and is “cleared” to be shared with anyone outside the VDE.

Only the cleared output should be used for publications or presentations to report your results. *Never* make the vetting requirements public in your publications.