



Virtual Pooled Registry (VPR) Project

Instructions For Creating the Annual VPR-CLS Registry Linkage File (8/31/20)

The Virtual Pooled Registry Cancer Linkage System (VPR-CLS) facilitates streamlined linkages between research studies and multiple cancer registries. Last year, 38 registries participated in VPR pilot test linkages and at least 8 additional registries are slated to come on board this year.

Each year as part of the NAACCR Call for Data, U.S. registries participating in the VPR initiative create a single, standardized VPR-CLS registry linkage file. This file will be used for every VPR linkage that occurs during a calendar year and serves as a static file from which researcher-requested data items will be selected. All VPR linkages occur behind each individual registry's firewall. **IMPORTANT NOTE:** We request that all U.S. registries, regardless of VPR participation, create the annual VPR-CLS linkage file to keep on file in the event that there are separate initiatives that utilize the VPR methodology in the upcoming year, similar to this year's Camp Lejeune Incidence Study.

Timeframe for File Creation: Registries will prepare the VPR-CLS registry linkage file as part of the annual NAACCR Call for Data process. The VPR-CLS files should be created after 12-month data (cases diagnosed in 2019) have been submitted to NAACCR, NPCR and/or SEER, but **no later than January 31, 2021**. Producing the VPR-CLS file immediately following the Call for Data ensures that all registry files contain comparable timeframes for follow-up and the most complete, high quality, edited data for linkage and data release. **NOTE: Unlike the Call for Data submission, the VPR-CLS registry linkage file contains patient identifiers and will remain at the cancer registry.**

Case Selection Criteria: Cancer registries will extract NAACCR Version 18 Type C records for all state resident cases. It is essential that all Type C data items are populated, including patient identifiers. Registries are encouraged to include cases from as far back as possible. At a minimum, include cases from inception year to 2019. The extract should include all reportable cases routinely submitted in the NAACCR data submission (e.g., in situ, invasive, and benign brain/CNS-prior to 2004 if collected) and include VA cases, DCOs, and cases received through inter-state data exchange.

Non-releasable cases: Non-releasable cases: Some of the above noted cases, particularly VA cases, may need to be removed prior to releasing individual-level data to the study investigators. If a registry has a field for non-releasable cases, it can be placed into the State/Requestor Items [NAACCR Item #2220] when the VPR-CLS linkage file is created. It is important that registries be able to identify and provide information about the cases that are unable to be released. Providing a summary of non-releasable cases is critical to inform the investigators of the potential for selection bias and to properly interpret study results.

NAACCR Prep: There is a specific configuration file for NAACCR*Prep to create the VPR-CLS file. This configuration generates additional fields as needed but does not delete the confidential information, first/last/middle name/SSN/etc, required to conduct a linkage.

Steps for Creating the VPR-CLS Linkage File

1. Create a fully populated Type C NAACCR layout file using the Case Selection Criteria listed above. This file can be created as a flat file or .XML.
 - a. If a registry has a field for non-releasable cases, it can be placed into the State/Requestor Items [NAACCR Item #2220] when the VPR-CLS linkage file is created. No other information should be placed the State/Requestor Items.
2. Download NAACCR Prep and the VPR-CLS NAACCR Prep configuration file from NAACCR Call for Data page.
 - a. The VPR-CLS NAACCR Prep configuration file that you download from the Call for Data page is configured in such a way as to allow all cancers diagnosed from 1973-2019 to be included in the output file. If you need to modify this range skip ahead to page 4, follow the instructions outlined on that page, and then return here.
3. Open NAACCR Prep. The Execution tab should be selected by default (see screenshots on pg. 3).
4. Press the button labeled “Specify Input”. Browse to the Type C file created in Step 1, select the file, then press the “Open” button.
5. Press the button labeled “Add Job”. The Add Job dialog box will appear.
6. Provide the location of the NAACCR*Prep configuration file that was created specifically for the VPR-CLS (either the default one or the one you modified in step 2a).
7. Specify where the output file should be written. Use the following naming convention to identify your state and diagnosis years (dates may vary by registry) when entering the name for the output file: ‘**XX(state)19952019**’. For example, if Idaho includes data from 1975 to 2019, the file will be named ‘ID19752019’.
8. By now, several drop-down boxes should be enabled at the bottom of the dialog.
 - a. Suppress County Fields (make sure “No” is selected)
 - b. Suppress Census Tracts (make sure “No” is selected)
 - c. Suppress Days on Dates (make sure “No” is selected)
 - d. Output File Type (make sure “XML” is selected)
9. Press the OK button to close the dialog. The job will be added to the execution queue.
10. Press the “Export Data” button.
11. The zip file that is output by NAACCR*Prep contains several files. We are only interested in the data file, which is the file with the .xml.gz extension. Extract the .xml.gz file from the zip file. The other files, and the zip file/container, can be discarded.
12. Review the output data file (the file with the .xml.gz extension) to verify that all Type C data items, including patient identifiers, are populated in the file. File*Pro (available on <https://seer.cancer.gov/tools/filepro/>) or other text editing tools can be used to view the file content.
13. Keep a copy of the file created in Step 12. This file will be used in future VPR-CLS linkages and data release activities.

NAACCR* 10.08.27

Execution Configuration

Specify Input **4** Specify the input file location

Add Job **5** Add one or more jobs to the queue

ID	Configuration File	Output File	Options	Status

Execution Log

Information will be displayed here when the export process begins.

10 Export Data

Cancel Export

Add Job

Provide the location of the NAACCR*Prep configuration file **6**

Specify where the output file should be written **7**

Job Options

Suppress Counties **8A** N/A Suppress Days on Dates **8C** N/A

Suppress Census Tracts **8B** N/A Suppress U.S. Fields N/A

NHIA Option N/A Suppress Canadian Fields N/A

Output File Type **8D** N/A

9 OK Cancel

Steps for Modifying the VPR-CLS NAACCR*Prep Configuration File

1. Open NAACCR*Prep. Select the Configuration tab.
2. Press the button labeled “Load Config”. Browse to the location of the VPR-CLS NAACCR*Prep Configuration File you downloaded from the NAACCR Call for Data website, select the file, then press the “Open” button. The configuration file will load.
3. At the bottom of the screen you should see 2 drop-down boxes representing the start and end points of the diagnosis year range. They will be set to 1973-2019.
4. Change the starting point to any year after 1973, but **no later than** the year of your registry’s inception.
5. Do not change any of the other configuration settings.
6. Press the “Save As...” button. Include the diagnosis year range that you chose in the file name (e.g. naacccr-vpr-cls-1995-2019.prep).
7. Return to step 3 on page 2.

NAACCR*Prep v2020.08.27

Execution Configuration

Fields to Retain (789) Fields to Calculate (66)

Move the fields you would like to retain to the list on the right

Search by Item #

Column #	Item #	Item Name
697-698	9445	AYA Site Recode
6155-6179	7010	Path Reporting Fac ID 1
6180-6199	7090	Path Report Number 1
6200-6213	7320	Path Date Spec Collect 1
6214-6215	7480	Path Report Type 1
6216-6240	7190	Path Ordering Fac No 1
6241-6260	7100	Path Order Phys Lic No 1
6261-6285	7011	Path Reporting Fac ID 2
6286-6305	7091	Path Report Number 2
6306-6319	7321	Path Date Spec Collect 2
6320-6321	7481	Path Report Type 2
6322-6346	7191	Path Ordering Fac No 2
6347-6366	7101	Path Order Phys Lic No 2

Column #	Item #	Item Name
1	10	Record Type
2	30	Registry Type
3-16	37	Reserved 00
17-19	50	NAACCR Record Version
20-29	45	NPI-Registry ID
30-39	40	Registry ID
40-41	60	Tumor Record Number
42-49	20	Patient ID Number
50-57	21	Patient System ID-Hosp
58-73	370	Reserved 01
74-123	70	Addr at DX-City
124-125	80	Addr at DX-State
126-134	100	Addr at DX-Postal Code

Configuration Options

Submission Type VPR Record Type Confidential DX Years 1973 2019 Survival Cutoff 2018

C:\Users\showew\Downloads\naacccr-vpr-cls-1973-2019.prep Last Modified: Monday, Aug 24, 2020, 01:38:52 PM