Virtual Pooled Registry (VPR) Project
Instructions For Creating the Annual VPR-CLS Registry Linkage File for December 2022 Submission

The Virtual Pooled Registry Cancer Linkage System (VPR-CLS) facilitates streamlined linkages between research studies and participating U.S. registries. Each year as part of the NAACCR Call for Data, VPR registries create a single, standardized VPR-CLS registry linkage file. This file contains PII, is 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. **IMPORTANT NOTE: <u>We request that all U.S. registries, regardless of VPR participation, create the annual VPR-CLS linkage file</u>** to keep on file in case there are separate initiatives that utilize the VPR methodology in the upcoming year, similar to the Camp Lejeune Incidence Study.

<u>Timeframe for File Creation:</u> 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 2021) have been submitted to NAACCR, NPCR and/or SEER, but **no later than January 31, 2023**. Producing the VPR-CLS file immediately following the Call for Data ensures that all registry files contain comparable timeframes for follow-up and include 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.**

<u>Case Selection Criteria:</u> Cancer registries will extract NAACCR Version 22 Type C records for all state resident cases. It is essential that all Type C data items are populated, including patient identifiers. At a minimum, include cases from inception year to 2021; however, registries are encouraged to include cases from as far back as possible. 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.

<u>Non-releasable cases:</u> Some of the above noted cases, particularly VA cases, may need to be removed prior to releasing individual-level data to the study investigators. It is important that registries be able to identify these cases and provide the researcher with summary of excluded cases. In previous years, if a registry had a field for non-releasable cases, it was placed in the State/Requestor Items [NAACCR Item #2220]; however, this field no longer exists as of v22 and the new fields for flagging these cases will not be implemented until v23. As a temporary workaround for this year, we are asking registries to place the information they previously placed into the State/Requestor items field into a new tumor level user-defined field we are calling "restrictedRelease". The XML-definition that you should use for this purpose has been provided below. If you are unable to create and populate the user-defined field as described, then your registry will need to manually identify and exclude non-releasable cases for each VPR study.
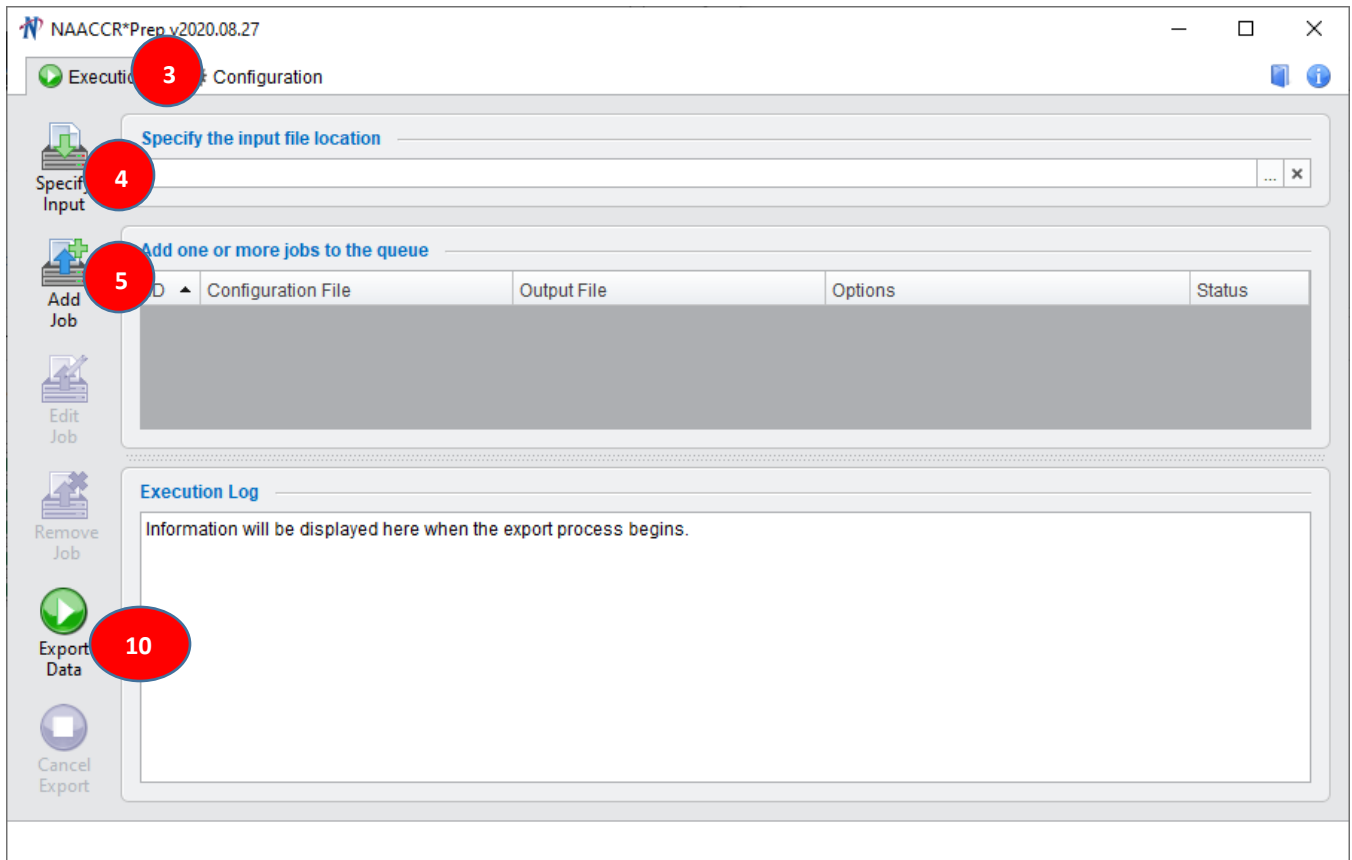
```
<ItemDef naaccrId="restrictedRelease"
      naaccrNum="9700"
      naaccrName="Restricted Release"
      length="1000"
      recordTypes="A,M,C,I"
      parentXmlElement="Tumor"/>
```

NOTE: Registries participating in the NAACCR Inter-Registry Data Exchange Agreement (Revised July 2021) allow the receiving registry to release inter-state-exchange-only cases for VPR studies unless otherwise stated in a registry addendum.

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 <u>fully populated</u> Type C NAACCR layout file using the Case Selection Criteria listed above. This file must be created using the .XML format.
   a. If a registry has a field for non-releasable cases, it can be placed into the Restricted Release field [user defined, #9700] when the VPR-CLS linkage file is created.
2. Download NAACCR Prep and the VPR-CLS NAACCR Prep configuration file from NAACCR Call for Data page.
3. Open NAACCR Prep.
4. Press the button labeled "Specify Input". Browse to the Type C XML 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 for the VPR-CLS.
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)YYYY(inception year)2021**'. For example, if Idaho includes data from 1975 to 2021, the file will be named 'ID19752021'.
8. By now, several drop-down boxes should be visible at the bottom of the dialog but only three are relevant (the remaining will be set to "N/A" and disabled)
   a. Inception Year (VPR): This field defaults to 1973, change this to your registry's inception year or the first year of high quality/complete data that you want to include in VPR linkages.
   b. NHIA Option: Choose "Option 1", "Option 2", or "All records".
   c. 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 and the log file. The data file has a .xml.gz extension. The log file has a .log extension. Extract both the .xml.gz file and the .log 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 .xml.gz file created in Step 11. This file includes PII and will be retained at your registry for future use in VPR-CLS linkages and data release activities.
14. Upload the .job.log file (some users may see just a .job file) created in Step 11 to the NAACCR Call for Data submission site under the Submission Forms section.

NAACCR*Prep v2020.08.27

Execution **3** Configuration

Specify the input file location

4

Specify Input

Add one or more jobs to the queue

5

D ▲	Configuration File	Output File	Options	Status

Add Job

Edit Job

Execution Log

Remove Job

Information will be displayed here when the export process begins.

Export Data **10**

Cancel Export