

Creating the Annual VPR-CLS Linkage Files

CASTINE CLERKIN
VPR PROGRAM MANAGER
NAACCR CALL FOR DATA WEBINAR
8/25/25

What's New?

- Starting in January 2026, the primary method for performing VPR linkages will be Privacy Preserving Record Linkage (PPRL), which leverages hash tokens rather than PII.
- IMS will perform automated VPR linkages for registries that have approval to upload a hashed registry file to the CFD portal.
- VPR Linkage Files will be created by December 31st.

Why PPRL?

- Enhances security of the VPR process
- Allows studies to use the VPR to link with registries if they cannot release a file with PII.
- Reduces the workload for VPR registries
- Facilitates quicker Phase I linkage results

Two Options for PPRL Linkage

- IMS, VPR Honest Broker, will run PPRL linkage on the registry's behalf if a registry has internal approval to upload a hashed file during the CFD submission process and can sign the required PPRL DUA.

OR

- Registry will run PPRL linkage behind their firewall if permission has not been granted to send a PPRL file to IMS.

Optional PII Linkage

- In rare situations (small cohort, rare cancer, low completeness of linkage variables), studies may qualify for PII linkage, like previous VPR linkages.
- Allows manual review of uncertain matches, which is not possible with PPRL.

VPR Linkage Files

PPRL Linkage File

- Small file that contains:
 - Registry ID [Item #40]
 - Patient ID Number [Item #20]
 - Hash tokens
 - Metadata fields (used to validate hash tokens)
- No PII included in this file
- Uploaded to the CFD Portal if IMS allowed to perform VPR linkages on behalf of the registry and registry can sign PPRL DUA
- Only used by IMS for automated linkages

PII Linkage File

- Large file that contains:
 - All NAACCR incidence variables, including PII
 - Hash tokens
 - Metadata fields (to validate hash tokens)
- Multiple uses:
 - Used for PPRL linkage when registry unable to send to upload file to IMS
 - Used by all registries for PII linkage
 - Static file used to export researcher-requested data items
- **Never uploaded to CFD Portal**

VPR File Creation

- Created after 24-month data submitted to NAACCR, NPCR, and/or SEER
- Created no later than December 31, 2025
- Request that all U.S. registries create both files, regardless of VPR participation
- Both files should be retained by registries for potential use
- Used for all VPR linkages that occur in 2026

Case Selection Criteria

- NAACCR Version 25 Type C records for all state resident cases with all data items populated, including identifiers
 - Do NOT mask or blank out the day portion of data fields. These are necessary for PII linkages.
- Cases from registry inception year to 2023 (24-month data).
 - Do **NOT** include cases from 2024 (12-month data)
- 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.
 - NAACCR Inter-Registry Data Exchange Agreement (Revised July 2021) allows the receiving registry to release inter-state-exchange-only cases for VPR studies unless specified in a registry addendum.

General Steps (detailed instructions on CFD Portal)

- Extract fully-populated NAACCR Version 25 Type C XML records
 - Non-releasable information should be coded to the two new NAACCR variables implemented in 2023: #1854 Patient No Contact Flag and #1856 Reporting Facility Restriction Flag.
 - All registries should populate Tumor Record Number [Item #60], which is a static identifier unique to each tumor belonging to a patient.
- Download NAACCR Prep and NAACCR Prep VPR Configuration File from NAACCR CFD site

General Steps (detailed instructions on CFD Portal)

- Follow detailed VPR file creation instructions for using NAACCR Prep to create Zip file of all necessary files.
 - File naming convention: 'XX(state)YYYY(inception year)2023'
- NAACCR Prep will create both the PPRL and PII Linkage Files, along with the XML user dictionaries, and .log file
- A 'salt' is used to create hash tokens and will be sent via secure email from IMS closer to December.

General Steps (detailed instructions on CFD Portal)

- Extract the PPRL and PII Linkage Files, XML user dictionaries, and log file
- Review PPRL and PII Linkage Files
- All registries upload '.log' file to CFD Submission Forms Section
 - This file is included in SEER*DMS data submission extract package
- If IMS performing PPRL linkage, upload PPRL Linkage File and initiate signing of VPR-CLS PPRL DUA via DocuSign under the DAA and Consent section
- Delete the salt file

Thank you!

VPR File Creation Instructions located
on NAACCR CFD Portal Documents Tab:

<https://www.naaccr.org/call-for-data/>



Contact: Castine Clerkin

cclerkin@naaccr.org

