Virtual Pooled Registry Cancer Linkage System (VPR-CLS) Fact Sheet

(Updated February 14, 2025)

The Virtual Pooled Registry – Cancer Linkage System (VPR-CLS) is an online service designed to:

- efficiently connect researchers performing **minimal risk linkage studies** with multiple U.S. population-based cancer registries.
- perform linkages between a study cohort file and cancer registry data files using standard linkage software and consistent matching algorithms.
- provide initial aggregate match count results to researchers.
- streamline the process of applying for release of individual-level data on matched cases.

Coordinated by the North American Association of Cancer Registries (NAACCR) with funding from the National Cancer Institute (NCI), the VPR-CLS provides a single location to facilitate minimal risk linkages between studies with an existing cohort and 47 U.S. registries representing approximately 96% of the U.S. population plus Puerto Rico. These registries collect high quality, complete, population-based cancer information, including patient demographics, cancer type, stage, treatment, and follow up. By providing a single point of access and streamlined processes to enable these linkages, the VPR-CLS significantly reduces the level of effort researchers must dedicate to the linkage, application, and approval process across registries.

The technology for the VPR-CLS has been developed by Information Management Services, Inc. (IMS), which also serves as the third-party honest broker. IMS has more than 47 years of information technology and clinical trials experience and employs a team of 250 computer and biomedical professionals located in the Washington Metropolitan Area. Long-term clients include the National Cancer Institute, the Centers for Disease Control and Prevention, the Food and Drug Administration, pharmaceutical companies, medical device companies, and other biomedical research organizations.

VPR-CLS Availability for Linkage with U.S. Registries

In February 2022, the VPR-CLS was officially launched. The VPR-CLS has facilitated 28 linkages to date. Interested researchers may contact the VPR Administrator (VPRAdmin@naaccr.org) to learn more about performing a linkage through the VPR-CLS.

The VPR-CLS is focused on minimal risk linkage studies. In order to utilize the VPR-CLS the following requirements must be met:

1. Study has (or will have) an existing cohort that meets the following criteria:
    - At least 75% of records contain 9-digit SSN, DOB and either First Name OR Last Name; OR
    - At least 75% of records contain First Name, Last Name, DOB, Gender, and at least one of the following: 4-digit SSN, Phone #, OR Full Address (excluding P.O. Box)
2. A current IRB-approval, IRB exempt determination, or documentation of Not Human Subjects
3. A protocol that includes linkage with cancer registries; and

4. A study consent form that includes linkage with cancer registries OR a specific waiver of informed consent to link with registries.

Resources to Streamline the Data Release Application and Review Process

The VPR-CLS streamlines the process of applying for release of individual-level data on matched cases by offering optional use of the following resources:

1. Templated Forms: NAACCR led efforts to create the following templated forms that can be used in lieu of the state-specific forms:

    a. Templated IRB/Registry Application (TIRA): The TIRA is a standard application that registries and IRBs use in lieu of their normal applications. The TIRA has been adopted by over 85% of the VPR-participating registries for all of their review process. The TIRA is filled out once during the VPR linkage process thereby significantly reducing the number of individual applications the researcher must complete.

    b. VPR Templated Data Use Agreement (VPR DUA): The VPR Templated DUA is a common agreement designed to be used in lieu of the individual registry DUAs. The VPR DUA has been adopted by over 50% of the VPR-participating registries to date, thereby minimizing the number of separate DUAs and ensuring consistency in the terms and conditions.

2. Central IRB: NCI has contracted with the Biomedical Research Alliance of New York (BRANY) to serve as a Central IRB (CIRB) for review of VPR linkage studies. For states that accept a CIRB for multi-site minimal risk linkage studies coming through the VPR-CLS, BRANY will perform the IRB review in lieu of state/local IRBs within a weeks' time. Among the 25 VPR registries that have an IRB, BRANY has a reliance agreement with seventeen (68%) of them, thereby improving the efficiency and consistency of IRB review.

Overview of VPR-CLS Workflow

All VPR-CLS linkage requests proceed in the following two phases:

- **Phase I** supports a standardized linkage with participating registries and provision of aggregate match counts (by registry and diagnosis year), allowing the researcher to prioritize registries from which to request additional record-level data in Phase II. Phase I includes a web-based application, secure data transfer protocols between researchers, IMS and registries, and use of a single record linkage software (Match*Pro) optimized for linkages between cancer registries and research cohorts. The Phase I web-based application also serves as a standard application used by nearly all registries for their Phase II registry/IRB application process.

- **Phase II** streamlines the process of applying to registries and/or their IRBs for release of individual-level cancer data for matched cases identified during Phase I. The system includes use of the CIRB, the TIRA, the VPR DUA, and a robust tracking system to monitor the study status across registries.

The Phase I match counts allow the researcher to review the volume of matches in each registry and make an informed decision about which registries to select for Phase II application for release of individual-level data. A detailed description of the Phase I and Phase II workflow is provided below.

Phase I: *Application to use the VPR-CLS to link with registries behind their firewall and receive aggregate match counts only (no state IRB or registry review needed).*

Anticipated Timeline: *Researchers can expect the application review process to take 2-4 weeks. After approval and upload of a validated, edited study file, the researcher can expect to receive registry match counts within 6-8 weeks depending on the list of studies in the queue for linkage.*

1. Researcher submits the online VPR-CLS application and supporting documents. Supporting documents include the current IRB determination, approved study protocol, consent form or waiver of consent, investigator's curriculum vitae, and signed DUA with IMS.
2. NAACCR reviews application and resolves any issues with researcher.
3. Research Review Committee (RRC), made up of seven representatives from cancer registries and key stakeholder organizations, reviews application and researcher notified of decision.
4. Researcher creates, edits, and uploads an encrypted study linkage file to the VPR-CLS in accordance with established file specifications and editing software.
5. Encrypted study file is validated and posted for pre-test linkage with select registries.
6. All registries perform linkage behind their firewalls using Match*Pro and standard linkage logic.
7. Registries create and upload an aggregate match count report to the VPR-CLS that includes the number of high quality and uncertain matches by diagnosis year (no patient records).
8. VPR-CLS reads the reports and presents researcher with the match counts.

Phase II: *Application for release of individual-level data on matched cases identified during Phase I.*

Anticipated Timeline: *IRB/Registry approval and release of data will vary based on the review process in each state and whether the TIRA (Templated IRB/Registry Application) and CIRB can be utilized.*

9. During Phase I, Researcher already submitted the TIRA and supporting documents for NAACCR review and once all registries have uploaded their match counts, the request proceeds to the next step.
10. Researcher reviews registry match counts and their adoption of VPR-CLS efficiencies (TIRA, VPR DUA, and CIRB) to inform selection of registries for Phase II, thereby finalizing the TIRA.
11. BRANY CIRB reviews the TIRA and enters approval for the relying registry IRBs in the VPR-CLS.
12. Researcher completes the remaining application materials and monitors the status of the request by leveraging the VPR-CLS list of additional required forms and agreements, interactive tracking system, and automated notifications and reminders.
13. Registries and local/state IRBs, as appropriate, review the study and sign agreements.

14. Upon approval and full execution of agreements, registries create a file of individual-level data for matched cases, including the Cohort ID and requested registry variables.

15. Registries transmit the data file directly to the researcher through a secure site, independent of the VPR-CLS, as specified by either the researcher or the registry.

Linkage Methodology

All VPR-CLS linkages are performed using the record linkage software, Match*Pro, developed by IMS. Match*Pro conducts probabilistic linkage based on the Fellegi and Sunter model. The following variables are used, as available, to link the study file with the registry file: First name, middle name, last name, maiden name, date of birth, social security number, telephone number, gender, and street address. After probabilistically identifying potential matches, deterministic filters classify each linked pair as a match, non-match, or uncertain.

Data Security and Protections

The VPR-CLS provides a secure, web-based portal through which researchers submit an application to use the system to link with registries. A DUA is signed between the researcher and IMS before the researcher uploads the study file containing patient identifiers. The website uses Transport Layer Security, ensuring that communication and files transferred between a client and the IMS server are securely encrypted. All study data files are encrypted by the researcher (using Match*Pro) prior to uploading them to the VPR-CLS. All files uploaded to the VPR-CLS are first scanned for viruses and then stored on a secure server behind the IMS firewall. Only authorized IMS staff can access the study files provided by researchers and all IMS staff have been trained in the handling of files that contain personal identifying information.

Once an uploaded encrypted study file has been validated by IMS, it is posted for secure download by an authorized liaison from each of the participating registries. All registry liaisons are authenticated and verified by IMS prior to receiving access to the VPR-CLS. In addition, each registry has confirmed compliance with a list of common security protections. Study files are used solely for data linkage by the participating registries. Each registry will perform the linkage behind their firewall. After the match count report is uploaded, the study file is no longer accessible for download from the VPR, and the registry is prompted to delete the study file and provide confirmation of destruction. After all data is sent to the researcher, IMS deletes the study file. The Match*Pro results file produced by each registry performing linkage contains the linkage results records. Those files are retained by the registries until the researcher requests to delete them.

Additional References

For more information on the benefits of linking with registries through the VPR-CLS, please refer to the following manuscripts:

- Virtual Pooled Registry-Cancer Linkage System: An improved method for ascertaining cancer diagnoses (J Natl Cancer Inst Monogr. 2024 Aug 1;2024(65):191-197)

- Ascertainment of Incident Cancer by US Population-Based Cancer Registries Versus Self-Reports and Death Certificates in a Nationwide Cohort Study, the US Radiologic Technologists Study (Am J Epidemiol. 2022 Jul 22;191(12):2075–2083)