General Instructions for NAACCR Call for Data
Deadline: Midnight Eastern Time Tuesday December 3, 2024

For questions about the NAACCR Call for Data, contact:

Recinda Sherman, MPH PhD CTR, Program Manager of Data Use and Research
Phone: (217) 698-0800 ext. 6; rsherman@naaccr.org

All Call for Data materials are located on our website:
https://www.naaccr.org/call-for-data/

****
STEP 1: Log in to NAACCR CFD System

Log-in to the Call for Data (CFD) Portal from the **Submit Tab**.

Click the long blue button "*Submit Call for Data, Complete CaRI database, Access Data Assessment.*"

- Prior to November 1, 2024, users will only be able to access CaRI submission section, their Data Assessment section, and only view their older submissions.
- NAACCR Dec 2024 CFD Portal will open no later than November 1, 2024.

Please ensure registry staff have proper access now. **Do not wait until you are submitting your file to log in**. Really. DO NOT WAIT.

If you have access issues, please contact Recinda Sherman rsherman@naaccr.org well in advance of the CFD deadline.

NOTE: Also ensure you are on the NAACCR **Call for Data listserv** to receive any updates throughout the CFD Process.

*Please contact Tricia Kulmacz tkulmacz@naaccr.org to check your status or add yourself to the **Call for Data listerv**.*

STEP 2: Duplicate Assessment

TUMOR- LEVEL MATCH*PRO PROTOCOL

Match*Pro, instructions, and all resources to apply the deduplication protocol are available on the CFD **Tools Tab.** It is important to run the protocol using the newest version of **Match*Pro 2.5.** This version will reduce the number of false positives for the patient-level deduplication and, subsequently, reduce the work on registry staff.

New this year: ***However,*** for this year only, the archive will not work properly on your tumor-level false positives. Please review the documentation BEFORE

downloading v2.5 and BEFORE starting your tumor-level deduplication: [Match*Pro Archive Issue](#).

*For technical questions about Match*Pro Duplicate Protocol, please contact Will Howe at howew@imsweb.com.*

The Match*Pro duplicate protocol is required for Certification and CiNA. We are in Year 2 of a 3-year roll-out for the updated duplicate protocol.

We have the following requirements for this year's CFD:

- Required: 100% Patient-level deduplication resolution on full submission (1995-2022)
- Required: Tumor-level match run on all cases diagnosed 2007-2022
- Gold = 2007+ with <.1% duplicates for all years of tumor-level resolution (2007-2022)
- Silver = 5 years of tumor-level resolution (2018-2022) with <.1% duplicates

Protip: You may consolidate identified duplicates and re-run Match*Pro as many times as you wish.

On the *General Information Form* in the **CFD Portal**, you will be asked to attest to the patient-level 100% resolution. NAACCR will run the tumor-level Match*Pro deduplication protocol on your submitted file to assess your tumor-level deduplication rate. Therefore, you are required to upload your final archive file along with your submission file. Detailed information about the archive files are found in the *NAACCR-Tumor-Deduplication-2024-CFD-MatchPro-Instructions* located in the zipped file on the **Tool Tab**.

You will also need to submit DAY of diagnosis. Submitting DAY ensures that we will be able to replicate the deduplication and resolution performed at your registry—and accurately assess your duplicate rate.

In NAACCR*Prep, you will have the following options for suppression:

Suppression Option	Action	Use
No – don't suppress day on any dates. NOTE: this is the default. New this year: This is the default for the US Only Supplemental submissions (NCCR, VPR, and NPCR)	no suppression on DAY data	Have option to include time to tx data for evaluation in CFD Portal Questions.
Option 1 – suppresses day on all dates EXCEPT for date of DX and dates of TX. New this year: This is the default for the CiNA Submissions (Certification and CiNA Research)	DAY of diagnosis, DAY of treatment ONLY submitted	Have option to include time to tx data for evaluation in CFD Portal Questions.
Option 2 – suppresses day on all dates EXCEPT for date of DX.	Submit DAY of dx ONLY submitted	DAY data is only used for purpose of Certification
Option 3 – suppresses day on all dates	No DAY data is submitted to NAACCR	More duplicates will be identified in your file.

Dates of Treatment, DAY submitted ONLY if selected:
NAACCR Item #1200   RX Date Surgery
NAACCR Item #1210   RX Date Radiation
NAACCR Item #1220   RX Date Chemo
NAACCR Item #1230   RX Date Hormone
NAACCR Item #1240   RX Date BRM

NAACCR Item #1250 RX Date Other
NAACCR Item #1260 Date Initial RX SEER
NAACCR Item #1270 Date 1st Crs RX CoC
NAACCR Item #3170 RX Date Mst Defn Srg

In the **CFD Portal**, you will confirm which suppression option you choose as well as give permission on how you will like the DAY of diagnosis information to be used—for the purpose of Certification only or for Certification and the to create "time to treatment" variables for evaluation purposes only (not to be released to outside researchers).

STEP 3: Create Submission File(s)

Create a single data file with all reportable cancer incidence cases for residents of your coverage area, who were diagnosed from 1995 through 2022. *The 2023 data are also being requested to be used for evaluation purposes only, including delay modeling and other assessments.* ***We strongly encourage you to submit the 2023 data, but this submission does not impact Certification or inclusion in CiNA products***. The 2023 data may be included in the submission file or submitted as a separate file as late as January 31, 2025.

We ***strongly recommend*** a full, confidential record dataset be used to create your submission file as NAACCR*Prep will remove variables not requested. Please request that your vendor supports this option to ensure the smoothest CFD process possible. However, SEER*DMS users use the report EXT-02 to create all submission files, and do not need to use NAACCR*Prep or a confidential dataset.

Submission Files

Registries should create a file that includes the data items listed in the excel document *Certification and CiNA Call for Data Items* available on the CFD **Documents Tab**.

A column on the excel spreadsheet indicates "New this Year" data items, so double check these variables are in your initial file pull. Again, it is ***strongly recommended*** to pull a full, confidential record as your initial extract file (unless you are a SEER*DMS user and can use report EXT-02).

A column on the excel spreadsheet also indicates "Required to derive but not submitted." These are data items that ***must*** be included in the initial file submitted to NAACCR*Prep, as they are necessary to generate derived variables. However, they will be removed by NAACCR*Prep and ***will not*** be submitted to NAACCR for Certification or CiNA. Again, if you are not a SEER*DMS user, we ***strongly recommend*** (are you sensing a pattern?) a full, confidential record dataset be used to create your submission file to ensure all requested variables are included. NAACCR*Prep will remove all variables not on the *Certification and CiNA Call for Data Items* list AND remove all variables required to derive but not submit.

Registries also have the choice within NAACCR*Prep to further suppress some requested variables.

The submission file:

- Data submitted must adhere to the NAACCR standard definitions and codes. Refer to the Standards for Cancer Registries Volume II, version 24 (https://apps.naaccr.org/data-dictionary) for details.
- All cases as defined in the document *Data Selection Criteria* on the **Documents Tab** for each diagnostic year should be included on the file, regardless of whether they are reportable in only some of the years.
- Data must be submitted in NAACCR Record Layout Version 24, XML data exchange standard.
- The derived fields are described starting at row 199 *Certification and CiNA Call for Data Items* excel spreadsheet on the **Documents Tab**. The spreadsheet includes links to documentation.
- <mark>New this year:</mark> We are requesting the two new geocoding quality fields generated the new geocoder, Geocodio. We will collect them as non-standard data items until they become NAACCR items (hopefully with v26).
  - o If you are able, you may submit these variables in your main submission file. You will need to review the additional instructions *How to Include Non-Standard Data Items Instructions* available on the **Documents Tab**.
  - o Alternatively, you can submit these two variables as a csv file along with a patient/tumor ID. NAACCR will append these variables to your data after submission.

STEP 4: Run NHAPIIA (US Only)

U.S. registries should run the NAACCR Hispanic and Asian/Pacific Islander Identification Algorithm on all cases submitted. NHAPIIA is incorporated into NAACCR*Prep. Unless you are a SEER*DMS user, w*e recommend using the NAACCR*Prep utility to generate the NHAPIIA fields.*

To run NHAPIIA in NAACCR*Prep, include the NAACCR items 2230, 2240 and 2290 (first, middle and last name) in the NAACCR*Prep input file. These fields will be used to derive the variables but will be deleted from the NAACCR*Prep output/submission file. Again, for non SEER*DMS users, we ***strongly recommend*** (yup, this again!) a full, confidential record dataset be used to create your submission file to ensure all requested variables are included.

If NHAPIIA is run outside of NAACCR*Prep, it is recommended that registries use Option 1, the current default for NAACCR*Prep. This setting limits the Spanish surname portion of the algorithm in counties that are less than 5% Hispanic to cases coded as surname only (item 190=7) or unknown whether Hispanic (item 190=9).

As part of submission Step 9, registries enter setting they selected—if another option is used, the registry will need to provide a rationale on the General Information

Submission Form. If your registry software incorporates NHIA/NAPIIA, please specify the version on the General Information Submission Form.

Users can control whether any existing NHIA/NAPIIA values are overwritten by NAACCR*Prep. You have the option to only run NHAPIIA on blank cases and not overwrite values generated at the registry. More details available on the NAACCR*Prep instructions.

STEP 5: Run Edits

Edits Metafiles are available on the **Tools Tab**.

a) Run the edit set that matches your submission.

All Edit errors on data items reviewed for Certification or included in CiNA submission must be reconciled prior to submission. The edits are based on the version 24 EDITS metafile releases.

For questions about Edits, please contact Jim Hofferkamp at [*jhofferkamp@naaccr.org*](mailto:jhofferkamp@naaccr.org).

Type of Submission/Type Registry	Edit Set
NAACCR Certification Edits-US, 2022	**NAACCR Certification Edits**
NAACCR CiNA-Canada, 1995-2023	**NAACCR CiNA Edits-Canada**
NAACCR CiNA-US, 1995-2023	**NAACCR CiNA Edits-US if not submitting to NPCR**
NAACCR CiNA-US, 1995-2023 (excluding NPCR CFD Edits)	**NAACCR CiNA Edits-US if submitting to NPCR**

The edit sets use NAACCR standard definitions and are based on submission of all requested data items. **WARNING!** If you choose not to submit a data item, you may get errors because of blank fields. If you like, you can turn these edits off. Be aware that we run EDITS on your submitted files and use our results to determine Registry Certification status and fitness of the data for various uses. If the submission type requires a variable, then the variable must be 100% error-free.

Canadian registries use the Canadian Edit set regardless of which submission file type. US registries can save some time by running the NPCR edits first and then the NAACCR edits using *NAACCR CiNA-US, 1995-2023 (excluding NPCR CFD Edits)*. Running the edits this way avoids running the same edit multiple times. Registries that do not submit to NPCR must us the Canadian or full US edits set, as applicable.

b) Run Inter-Record Edits

Run the Inter-record edits ***after*** running the main Edits Metafile. Please resolve all errors.

Correct all errors prior to uploading the file to NAACCR*Prep. Registries are not required to re-run edits the NAACCR submission file generated by NAACCR*Prep through edits.

NOTE: *There is no combined NAACCR/NPCR edit set. All materials necessary for NPCR submission will be provided by NPCR. U.S. registries submitting data to both NPCR and NAACCR will need to use different edit sets for each agency's CFD.*
NOTE: *Only one error free edit run is necessary. Submission of an edit summary report is not required.*
NOTE: *The 2023 data is for evaluation purposes and assessment of continuing Covid-19 impact on cancer reporting only. The 2023 cases should be as clean as is practical but do not need to be error-free nor deduplicated at the tumor-level.*

STEP 6: Create Output File(s)

NAACCR*Prep and the configuration files will be available on the **Tools Tab** by the middle of October 2024.

The submission file should be free of edit errors prior to running the NAACCR*Prep utility (see steps 5).

*NOTE for SEER*DMS Users*: The SEER*DMS report EXT-02 should be used to create the NAACCR CFD file as well as the VPR-CLS file for SEER and non-SEER registries. EXT-02 has been updated to include a VPR.log file to accompany your submission file. NAACCR*Prep is not needed for NAACCR submissions for registries using EXT-02. **All required fields – including the calculated fields have been added to the extract.**

Protip: If your file does not get saved or gets corrupted, or if your VPR-CLD log file (Step 12) does not generate, save your output to somewhere other than a network drive.

NOTE: *This VPR-CLS file will not leave your registry. Any linkage conducted will be behind your registry firewall.*

The NAACCR*Prep program generates the derived variables and generates the submission file. Derived variables include the survival variables and derived grade for all registries—and the area-based social measures (e.g. RUCA, URIC, poverty codes, and Yost) and NHAPIIA for U.S. Registries. The program automatically strips specific fields from your submission file after the derived fields are developed.

The variables that are automatically stripped during processing are the name fields and day of last follow up. If will also automatically suppress US-specific fields for Canadian registries and Canada-specific fields for US registries. The EPHT fields are also suppressed for the NAACCR submission.

If your state requires, you also have the option to strip the following variables:

- Census tract 2010/2020 – data collected **only for evaluation and method development and will not be released** to researchers
- County at Diagnosis (all county fields) – data **are never released for research** without specific permission from your registry.
- Days of dates – **not recommended to strip** as duplicate rate will not be properly assessed (See Step 2).
- Cause of death – For general survival, the SEER recodes are used. COD data **are never released for research** without specific permission from your registry.

For NAACCR*Prep generate the derived fields, all the "Required to derive but not submitted" fields as indicated on *Certification and CiNA Call for Data Item* spreadsheet must be included in the input file including NAACCR items 110, 130, 2230, 2240, 2390. These variables **will never be included** in the output file submitted to NAACCR.

The *Certification and CiNA Call for Data Items* spreadsheet available on **Documents Tab** has additional information about the fields generated in NAACCR*Prep: the calculated survival variables, derived summary grade, and the geographic-based measures (U.S. only).

Geocoding Quality Variables:
Registries that use the batch match, on-line version of Geocodio can submit a modified csv file based on the csv file output by the geocoder. Just ensure the linking variables (PatientID/TumorID) and the 2 data quality variables (Status and Penalty Code) are included and delete <u>all</u> other variables. This limited csv file is then your submission file.

Some registries may have incorporated these variables into their registry and can include them in their regular submission file. If this is the case, you will need to review the additional instructions H*ow to Include Non-Standard Data Items Instructions* available on the **Documents Tab.**

<mark>New this year:</mark> Starting with v24, the two data quality variables that were only generated by the old NAACCR Geocoder (Micromatch Status and Penalty Code) are still requested using their NAACCR standard names and definitions (#86 Geocoding Quality Code, #87 Geocoding Quality Code Detail).

STEP 7: Name Output File(s)

Create **one** data file for your NAACCR submission. For example, if submitting 1995 through 2023 data, the output file should be labeled using the naming convention ***SS9523v24.xml*** where SS is the two letter postal code for your registry, followed by the years of data submitted, followed by the NAACCR version layout. For example, Wyoming submitting data in NAACCR version 24 for the years 1995-2023 would create a file named ***WY95223v24.xml***.

If you must submit multiple submission files, use the same naming convention.

Using the Wyoming example, a Certification only submission would be **WY2022v24.xml** and the 12-month data submission would be **WY2023v24.xml**.

NOTE: *NAACCR*Prep will create a zip file for you with the proper extensions.*

NOTE: If you are submitting a separate CSV file with the geocoding quality variables, Geocoding Accuracy Score and Type, follow the same naming convention. Using the Wyoming example, a separate geocoding quality variable submission would be **WY9522v23.cvs**.

STEP 8: Compress Data File(s)

NAACCR*Prep outputs a zipped file. If you choose not to use NAACCR*Prep or you are submitting a separate csv file with the geocoding quality variables, you must compress the data file with 7-Zip, WinZip, PKZip, or Gzip. Gzip is the preferred method for compression. Gzip executables for Microsoft Windows and other operating systems are available for free at http://www.gzip.org/#exe.

Protip: *If you wish to review your NAACCR*Prep output file prior to submission, NAACCR*Prep can create a csv file for review or you can use the resources available on the* **Tools Tab** *under "Convert Between XML and Flat File." You can also use File*Pro to view your submission file.*

STEP 9: Submit Submission File(s)

The CFD Submission Portal will be open no later than November 1, 2024 and is accessed from the **Submit Tab**.

- Click the long blue button "*Submit Call for Data, Complete CaRI Database, Access Data Assessment*" on the *Submit* Tab.
- Click "*Submit the 2024 Call for Data*" in the 2024 CFD Box
- Click on "*File Submission Forms*" complete the form and upload your file(s).
- Once this step is complete, this section will show as checked on the dashboard.

STEP 10: Complete Submission Forms

Complete all of the electronic submission forms in the CFD Submission Portal by **December 3, 2024**.

- From within the CFD Submission Portal, click on links, for example "*General Information*" or "*Process Verification & Vital Status Follow-Up*" to complete all relevant forms.
- As you complete this information, these sections will show as checked on the dashboard.

STEP 11: Review, sign, and submit the NAACCR Call for DAA and Consent Forms and mark COMPLETE

Review, sign, and submit the DAA and Consent Forms in the CFD Submission Portal **by December 3, 2024.**

- From within the CFD Submission Portal, click on link "*Sign DAA and Consent Forms*" to sign the forms via DocuSign. **DO NOT SIGN THE VERSION AVAILABLE FOR REFERENCE ON THE DOCUMENTS TAB.** These are for reference only and DocuSign contains the final versions.
- As you complete this information, this section will show as checked on the dashboard.
- Once all the forms are signed and submitted, and all the sections are checked, **please mark your call for data as COMPLETE**.

NOTE: *Registries have the option to e-sign using the industry standard, DocuSign, or download, sign, and upload all submission documents that require a signature. Follow the instructions on the Submission Site.* ***If you download, sign, and upload – please indicate "Signed on paper" on so that it can be appropriately tracked.***

NOTE: *Occasionally, a registry needs to make a small change to the DAA. This can be done by downloading the DAA from DocuSign.* ***If your registry is required to do this, please indicate the "DAA Has Modifications" so it can be appropriately tracked.***

NOTE: *If you submit 2023 data after your regular submission, you may need to contact NAACCR (*rsherman@naaccr.org*) to "Unlock" your Submission.*

Checklist of materials requiring signed approval:

- ✓ **NAACCR, Inc. Call for Data Assurances Agreement**
- ✓ **ACS Facts and Figures [US Only] – Active Consent**
- ✓ **CiNA County Maps [US Only] – Active Consent**
- ✓ **Mayo Catchment Area [US Only] – Active Consent**
- ✓ **Medullary Thyroid Cancer [US Only] – Active Consent**
- ✓ **CiNA Public Use Dataset—Passive Consent**
- ✓ **American Lung Association [US Only] – Passive Consent**
- ✓ **Delay Adjustment—Passive Consent**
- ✓ **Mapped County Modeled Data—Passive Consent**
- ✓ **The Affordable Care Act and Cancer Stage at Diagnosis [US Only] – Passive Consent**

STEP 12: Create VPR-CLS File (US Registries Only)

After all other Call for Data Certification & CiNA submission requirements are complete, please create your linkage file for the Virtual Pooled Registry Cancer Linkage System (VPR-CLS) for US registries only.

New this year: The VPR-CLD file is due **by December 31, 2024** (one month earlier than prior years).

Upload your log file created in NAACCR*Prep that is to be uploaded under the *VPR_CLS File Log* section of *Submission Forms*. **Only the log is to be uploaded to the portal. Your VPR-CLS file stays at your registry, behind your file wall.**

VPR-CLS is a voluntary NAACCR initiative to facilitate and streamline the process of linking research cohorts with multiple cancer registries. We ask that all registries create the VPR-CLS File at same time to ensure comparable timelines for follow-up. If you have any questions on the VPR-CLS file, please contact Castine Clerkin cclerkin@naaccr.org.

While this is a voluntary step, we request *all* U.S. Registries to create this file during the Call for Data regardless of participation in the VPR-CLS to support potential, national initiatives, such as the Camp Lejeune Cancer Incidence Study which linked with all U.S. states, DC, and PR—but was not part of the VPR-CLS.

Detailed instructions for creating the VPR-CLS linkage file are on the **Documents Tab**. Also please review the *How to Include Non-Standard Data Items Instructions* on the **Documents Tab** if you are including the "restrictedRelease" or "registryDefinedTumorID" variable in your file.

STEP 13: CaRI (Cancer Registry Information) Database

After other CFD requirements are complete, please log into the CFD Submission Portal on the **Submit Tab** to review, update, and/or enter the registry information for the CaRI database *by January 31, 2025*. This step is for both US and Canadian Registries.

This database captures essential information on cancer registry data. The system is populated by cancer registry staff and reviewed annually at a minimum. The data is publicly available in a queriable format here: https://www.naaccr.org/cari-database. If you have any questions, please contact Recinda Sherman rsherman@naaccr.org.

Once logged in to the CFD Portal, click "*Enter registry information into the CaRI Database*" in the top right box. Once your entry or update is complete, click on the "*Mark as reviewed*" button. To review and create specific queries of the data, go to the public site at: https://www.naaccr.org/cari-database.

NOTE: *Registries have input information in previous years. Those registries will only need to review and update if needed **but you still need** to click the "Mark as reviewed" blue button.*

STEP 14: Create National Childhood Cancer Registry (NCCR) Submission File (US, NCCR, Non-SEER Participating Registries Only)

NCCR participating registries who are also SEER registries will have IMS submit the

NCCR file on behalf of the registry. NCCR participating registries who are not SEER registries will create and submit their own submission file.

Detailed instructions for creating and submitting the NCCCR file will be available in the *NCCR Submission Instructions* file on the **Documents Tab**.