

NAACCR Call for Data (CFD) Instructional Webinar

Q&A

August 28, 2023

#	Question	Asker Name	Answer
1	Hi Recinda, Can registries have multiple users with access to the CFD portal? or will registries need to select 1 staff person?	Anonymous Attendee	live answered; Yes. Any user with Access Role in MyNAACCR for Proxy Form Submitter or Director will have access to the CFD Portal for submission. Any user associated with the registry in MyNAACCR will be able to access other areas of the portal, such as the Data Assessments. Registry Primary Contacts can assigned Access Roles for individual users, as well as the users themselves.
2	We need more information on why "day" is needed and why we are strongly urged to use a Confidential File. Both requests will require further discussion at our registry as to whether we can submit.	Jennifer Hafterson	live answered; also documented in the CFD instructions
3	match*pro archive file: what are the contents, anything identifiable?	Valerie Yoder	live answered; a hash code and patient ID—nothing identifiable. You will be able to review that file as well if you wish.
4	Why does a "log" need to be sent in to NAACCR for VPR-CLS? How does NAACCR use that log?	Anonymous Attendee	live answered; used by NAACCR to confirm VPR-CLS dataset has been created by January 31 st . It can also be used by the registries to document where that file is stored. There is no identifiable information in the log file.
5	Recinda - is there a custom restricted release flag specific to VPR (as opposed to the standard restricted release for ALL DoD/VA/Out-of-state)? Since VPR studies allow	Jenna Deniaud	Hi Jenna and others, yes. As last year, we are allowing a custom restricted release variable. All needed information is in the VPR-CLS and How to Handle Non-Standard Variables sets of instructions on the Documents Tab.

	release of many Out-of-state cases per the newest inter-registry data exchange agreement?		
6	Has any adjustment been made for the Race quality criteria since it is getting more difficult to get particularly from pathology and physician reporting?	Gary Levin	Nothing planned at this time. Race collection is a focus of NAACCR-- but for now, we are focused on better collected and processing. But we are gearing up to review the Certification requirements for potential update. So there might be changes at that time. The process of review and potential changes will involve the NAACCR community, and registries will write any recommendations for Board approval.
7	Hello! DCO followback: Can you elaborate on registries will be contacted with over 0.5%?	Vijay Medithi	Hi Vijay, the threshold is below 0.5%. If your DCO % is extremely low (<.05%), it may be that full case ascertainment from mortality files was not performed. I will be following back with any registry with <0.5% to better understand their death clearance process before determining if eligible for Certification. The slide had a typo and has since been fixed to reflect <i>less than</i> .
8	Do we still need to run IR Edits tool as well?	Tuyet Thieu	Yup! Both edits are on the Tools Tab. Jim will discuss more and the General Documentation discusses both.
9	What if we do not use the NAACCR geocoder due to internal restrictions.	Georgia Yee	Same as last year -- then nothing to submit :) However, please reach out to see if we can resolve any issues. We are moving to another NAACCR recognition focused on geospatial epidemiology Fit for Use. So a registry that does not process the geocoded cases standardly might be ineligible for that additional recognition. But this is in process--we don't know exactly how that recognition will sort out.
10	Thanks for the clarification. I read that as >0.5%	Vijay Medithi	The typo has been fixed in the slide uploaded to the CFD landing page.
11	If there a new one? Where can a new version of 2023 Inter Record Edits tool be found?>	Anonymous Attendee	All needed resources are on our Call for Data Landing page. The Edits, including the updated 8/23/23 IR Edits, are posted there.
12	Can we get a tumor deduplication configuration that does not use "day" so that if we choose not to submit "day", we can review the slew of false positives?	Jennifer Hafterson	live answered; the 15 th will be generically used—resulting in potentially many more cases to review.
13	What specific changes were made to Tumor/MPH library? Is a change log available?	Gary Levin	The change log can be found here: https://github.com/imsweb/mph/releases . These changes reflect an NCI update to the rules/application of the MPH rules.

14	The custom (user-specific) data item is only used if registries have not yet adopted the new standard restricted release flags [Items #1854 and #1856]. The new flag fields are designed to be used for the VPR if registries have adopted them (which is recommended).	Castine Clerkin	Thanks, Castine! Note: All these details can be referenced in the comprehensive VPR-CLS instructions. Thank you, Castine, for making detailed instructions available.
15	Other than the DOB change, were any other changes made in the Patient Deduplication matching logic?	Gary Levin	First name needs to be a fairly good match if SSN doesn't match exactly, regardless of how well everything else matches, in order for a pair to be classified as a high-quality match.
16	Does the Tumor deduplication do matching on all years or only what is required by NAACCR for certification?	Gary Levin	The tumor deduplication process looks at cases diagnosed in 2007 or later. For silver certification, registries are required to review and then consolidate matches and archive non-matches for all duplicates that contain one or more cases diagnosed in 2021. For gold certification, registries are required to review and consolidate/archive all duplicates that contain one or more tumors diagnosed between 2017 and 2021.
17	Are there instructions to create the Archive when Match*Pro is not used to clear matches?	Gary Levin	Instructions on how to do this are in Match*Pro. To access this function, go to Tools->Match Status Archive Manager and choose "Create a status archive from an external file" from the dropdown at the top. You can press the Question Mark icon next to the dropdown to bring up the help that explains how to use the function.
18	Just to clarify, CCRs need to perform the de-duplication process at the patient level and at the tumor level correct?	Lucinda Ham	Hi Lucinda, in order for the tumor-level deduplication to be accurate, full patient-level needs to be conducted first. Yes, long way to answer, yes, you are correct, and the patient-level should be done first.
19	Are the tumor hash codes exportable?	Gary Levin	Yes, after running the linkage you can press the "Export Data" button to bring up a dialog that will allow you to export a CSV containing any fields that you want, including the "cancerHash [calculated]" field. This allows users to generate an archive file without monitoring the archive within Match*Pro (for users that resolve and track false positives outside of Match*Pro).
20	Will - when asking how many were resolved (slide 12 i think?) - does resolved = true positives/duplicates which have been consolidated/fixd?	Jenna Deniaud	live answered; Will is using resolved to mean the identified duplicate case was consolidated (or deleted). However, for the purpose of calculating a duplicate rate, we also loosely consider the false

			positives that are marked as false positives in the archive as “resolved—” meaning they won’t be included in the duplicate rate.
21	Will there be a place in the NAACCR CFD portal to upload the Tumor Status archive file ?	Jason Jacob	Yes, as mentioned, you will upload your archive file along with your data file. All details for handling the archive file are in the Match*Pro instructions.
22	Since the tumor deduplication process is time consuming, we will need to start it while we are still processing data. So we will likely resolve and run it one more time before data submission. From which review do we grab the numbers for CFD page?	Ruth Li	live answered—send in the numbers from your final Match*Pro run.
23	So if we have 10 potential duplicates and determine 8 are duplicates and 2 are not...then only 8 are considered resolved? The 2 are really not duplicates...but don't consider them resolved?	Lyn Almon	Hi Lyn, so the terms are a bit confusing. Those two cases are false positives. Once your registry reviews and determines they are truly NOT duplicates, then you mark that in your archive and those 2 cases will be removed from your duplicate rate when we assess it. Any case that was a true duplicate and is "resolved" in the database (consolidated or deleted) will no longer show up when the protocol is run either a subsequent time at your registry or when we run on your file. You can think if it as the duplicates are "resolved" and the false positives are “documented.”
24	Since DX Date allows for partial dates then there will be times when month and day missing.	Gary Levin	We will need the full date of diagnosis in order to accurately assess the number of duplicate tumors. If month and day are missing, Match*Pro will assume a date of July 2 (the midpoint of the year), which may not be what you want and may lead to false positives or false negatives, so for that reason we would need the full date.

25	Our registrars (AB, Canada) don't use Match*Pro MP) to review the duplicated tumors. Our analysts use MP to run duplicated cases and export results and send them to our registrars for reviewing (as we do for IR Edits with override flag such as IR09). So, how can we do the same process for MP as we do for IR Edits without getting our registrars use MP to review the results directly on MP?	Anonymous Attendee	<p>When the analysts send the results to the registrars to review (assuming they use the Match*Pro export to CSV function to do so), make sure they include the File 1 ID and File 2 ID out of Match*Pro. These IDs consist of the patient ID, followed by a semi-colon, followed by a cancer hash value.</p> <p>The registrars should indicate which of the pairs sent to the were or were not duplicates. The analysts can use this information to update the match status value in the linkage results file and create the archive file.</p>
26	Ok...so there is a place to note those 2 in the archive as NOT duplicates?	Lyn Almon	That is the purpose of the archive—to mark the cases that you have reviewed and determined are NOT duplicates. The archive documents, tracks, and removes those that you indicate as false positives from having to be reviewed again (unless essential information is updated—then you would potentially need to review again).
27	Ok..thanks Recinda for clarifying	Lyn Almon	
28	Is there a way for NAACCR Prep to be able to export all the derived fields along with PID/Central Sequence or is this something that File Pro can do? We would like to update our database with these values.	Gary Levin	You can either convert your submission file to a csv file for manipulation (using File*Pro or other system of your choice). Item #20 and #380 are in the submission file, so you can combine. OR you can use File*Pro to create the fields. There are additional fields available in File*Pro--some we calculate on the back end or that we support for other purposes. Either method is fine.
29	This could be a huge burden depending on the number if false positives.	Gary Levin	You should be seeing fewer false positives now.