PUF FAQs

Administrative Questions

Q: I am not at a Commission on Cancer Accredited Facility. Can I request PUF data?
A: No, only investigators from CoC accredited facilities can apply for PUF data.

Q. Can I have another person complete the application for me if I am the Principal Investigator?
A: The application cannot be started without agreeing to the Terms of Agreement. Only the Principal Investigator can sign the Terms of Agreement as they are responsible for the data.

Q: Can I have a statistician or analyst at another facility analyze data for my PUF?
A: No. Per the Terms of Agreement, no person outside of your facility may have access to the data.

Q: Can I have access to NPI numbers, zip codes, etc.?
A: No, we cannot fulfill special requests for data not included in the PUF. The *Data Items* tab of the online data dictionary (http://ncdbpuf.facs.org/) contains information about every data item included in the PUF.

Q: Why aren't my login credentials working to access the PUF Application Manager?
A: You must be designated as a 'NCDB PUF Applicant' in CoC Datalinks in order to apply for a PUF. A cancer registrar at your facility can do this for you. If you have forgotten your log in credentials, please use the following link to request your login information:
https://web5.facs.org/Cancer/Account.mvc/EnterEmail

Q: Why isn't my password working to open my data files after I've downloaded the file?
A: This password isn't the same as the one you use to log into the PUF Application Manager or Datalinks. After you've logged into the PUF Application Manager, you will find a unique password for each of your current applications on the right-hand side of the screen.

Q. I would like to add investigators to my application after it has been approved. How would I do this?
A: Currently you cannot log in and change your application once it has been approved. To add investigators you will need to e-mail the biosketch and e-mail address for each new investigator to NCDB_PUF@facs.org. Please note that any investigator with access to the data needs to be included in your PUF application. If you are adding an analyst/statistician or the analyst/statistician is no longer at your facility you will need to update your application.

Q. I would like to change my research plan or add an additional research plan after it has been approved. How would I do this?
A: Currently you cannot log in and change your application once it has been approved. To add research questions you will need to send your revised plan to NCDB_PUF@facs.org. It you are making only minor changes you don't need to inform the NCDB.

Last updated 4/12/17

Data Questions

Q: Are NCDB data population-based?
A: No. The NCDB are hospital-based, not population-based. Please do not refer to the NCDB as population-based.

Q: Can I calculate incidence using NCDB data?
A: Since the NCDB data are not population-based, you cannot calculate incidence.

Q: Where can I find information on case counts for primary sites?
A: Case counts for primary sites, histology, and select variables, such as demographic and treatment items, may be accessed through the Public NCDB Benchmark Reports website: https://www.facs.org/quality%20programs/cancer/ncdb/publicaccess/.

Q. Can you provide data for a selected cancer site with frequencies for selected variables I am interested in?
A: No, the NCDB cannot run any special analyses for PUF applicants. You have access to the case counts by site and diagnosis year described in the previous question only.

Q: How can I investigate Head and Neck cancers or any other cancer site with an unknown primary?
A: Clinical T0 may be used for sites where the primary is unknown. The primary site code in your site specific cancer file indicates what the physician believed the primary site to be for clinical T0 cases. The ICD-O-3 code C80.9 is not a useful code and is not provided in the PUF.

Q: Do you have { } site or { } histology?
A: Check our site-histology groupings: http://ncdbpuf.facs.org/node/416. (Tip: If you are interested in Head and Neck sites, make sure that if you are planning on including Nose/Nasal Cavity/Middle Ear and Larynx, that you select those in the *Respiratory System* group in the site selection section of your PUF application. The rest of the Head and Neck sites are included in the *Head and Neck* group.)

Q: I am interested in analyzing pleural mesothelioma. Which site should I request?
A: You should request *Mesothelioma* in the site selection section of your PUF application. Mesothelioma histologies are not included in the *Pleura* site.

Q: I am interested in melanoma of sites other than skin, what files would I request?
A: You should select the primary sites of interest, with the understanding that some primary site codes may include histologies outside of the 8720-8790 histology range.

Q: I am interested in sarcomas not included in the Soft Tissue and Heart primary site files. How would I obtain these?
A: You will need to select the primary sites of interest. As above, when analyzing the data, you would then select only the sarcoma histology codes you are interested in. See https://www.iarc.fr/en/publications/pdfs-online/epi/sp155/ci5v8-chap4.pdf for more information about sarcoma histology and primary site codes.

Q: Does the PUF data include recurrence data?
A: No. Although the NCDB collects recurrence data, it is not included in the PUF because there is a large percentage of missing data. See the following article for more information: https://www.ncbi.nlm.nih.gov/pubmed/24504926.

Q: Are data on salvage therapy available in the PUF?
A: No. Only first course treatment is collected by the NCDB and included in the PUF. First course treatment is defined as all methods of treatment recorded in the treatment plan and administered to the patient before disease progression or recurrence.

Q: Why are there data missing for the Site Specific Factor I am analyzing for certain diagnosis years?
A: Site Specific Factors were first collected in 2004 as part of the Collaborative Stage Data Collection System. There have been changes in the Site Specific Factors (SSF) over time, with some SSF being added in 2010, some SSF that were revised, some that were only voluntary, and some that were discontinued. For example, breast cancer HER2 status was added in 2010 using SSF 8 through SSF 16, so there will be no data for these variables before this year. In some cases, registrars may have resubmitted older cases (before 2010), and added HER2 status for these patients. However, the data before 2010 will be sparse and should not be used. In other instances, the coding for a variable changed, such as for Prostate cancer Gleason Score.   SSF 5, Gleason's Primary Pattern and Secondary Pattern Value, and SSF 6 Gleason Score, were collected in 2004-2009. In 2010, two new SSF replaced these, SSF 7, Gleason's Primary Pattern and Secondary Pattern Values on Needle Core Biopsy/TURP, and SSF 8, Gleason Score on Needle Core Biopsy/TURP. If you are analyzing Gleason Score between 2004 and 2010 and later, you will need to use the SSF in use for the specified diagnosis years.   Another example using Anal cancer, SSF 1, HPV status, was a voluntary SSF item collected in 2010, but was discontinued in 2014.   In presenting your findings for the SSF, do not refer to data as missing for Site Specific Factors that were not collected in selected diagnosis years. Instead, indicate that data were not collected for those years.   For more detailed information about the Site Specific Factors, go to the PUF Data Dictionary Page at http://ncdbpuf.facs.org/node/367.

Q: Some of the cells in tables in my manuscript have fewer than 10 cases. Since it is prohibited by the Data Use Agreement to report cells with <10 cases, what should I do?
A: Here are some suggestions:
1.   Do not include the total n for any column so you cannot calculate the number in a particular cell that is suppressed.
2.   Delete the offending row entirely without indicating what the row represents anywhere,
3.   Combine rows so there will be more than 10 in each row.

Q: Does the PUF include cases treated at more than one hospital?
A: Yes. A varying proportion of cases by disease site will have treatment administered across programs, such as those receiving surgery at one hospital followed with adjuvant treatment at another. To simplify your analysis, the NCDB has selected the best analytic record from multiple case submissions across hospitals so that each unique case is represented as one row in the PUF file. All registries at CoC-accredited programs follow-up with patients and will report treatment as either "summary" (treatment at any hospital) or "at this reporting facility". These fields are listed under the treatment tab of the PUF data dictionary. Note that the facility-specific metrics contained in the PUF (such as location, type, and distance) only apply to the facility that submitted the case and administered treatment.

Q: Can I link cases of a patient with multiple primaries?
A: No. A case can be identified as having been diagnosed with multiple primaries by using the Sequence Number data item (http://ncdbpuf.facs.org/content/sequence-number), but the NCDB does not have any way of linking these primaries.

Q. I have a PUF data file from a previous application cycle. Can I combine it with a newer file?

A: No. In addition to a new diagnosis year of data being added each calendar year, existing cases may be updated with vital status, treatment status, or other follow-up information. Duplicates of cases may also be present if attempting to combine data files of the same cancer site from multiple versions of the PUF. Additionally, new random facility IDs are generated for each version of the PUF, so you would not be able to match facilities for analytic purposes.

Q: There are missing data for variables that should be aligned. For example, there are cases where AJCC Clinical T is missing but Clinical Stage Group is not. Or, Regional Nodes Examined is missing but Scope of Regional Lymph Node Surgery is not.

A: Facilities are responsible for maintaining the quality of their data, and each facility creates their own Quality Control Plan, per Standard 1.6 of the CoC Standards (see page 35). For more information about data quality checks, please consult the registrar at your facility. They also receive Completeness Reports from the NCDB detailing missing data after the data submission period each year. However, there may still be minor inconsistencies in some variables due to human error.

Q. Do you have any additional resources from which I can learn more about NCDB data?

A: For more information on the NCDB, you can read the following manuscript:
Boffa DJ, Rosen JE, Mallin K, Loomis A, Gay G, Palis B, Thoburn K, Gress D, McKellar DP, Shulman LN, Facktor MA, Winchester DP. Using the National Cancer Database for Outcomes Research: A Review. JAMA Oncol. 2017 Feb 23. doi: 10.1001/jamaoncol.2016.6905. PMID: 28241198.