

2023 NATIONAL AMBULATORY MEDICAL CARE SURVEY HEALTH CENTER (NAMCS HC) COMPONENT TECHNICAL DOCUMENTATION

For Public Use Data File



Division of Health Care Statistics
National Center for Health Statistics
August 2025

Overview Summary

This document provides detailed information and guidance for users of the 2023 National Ambulatory Medical Care Survey Health Center (NAMCS HC) Component public use data file. As a principal source of information on health care utilization in the United States, the NAMCS HC Component collects visit data from a nationally representative sample of U.S. federally qualified health centers (FQHCs) and FQHC look-alikes through electronic health record (EHR) data submission. The 2023 NAMCS HC Component is conducted by the National Center for Health Statistics (NCHS) and is a member of the National Health Care Surveys – a family of surveys which measure health care utilization across a variety of health care providers and settings.

Section 1 of this document includes information on the scope of the survey, the data sources, and the confidentiality protections related to the data. Section 2 contains details on the sampling process, data collection procedures, and weighting methodology used to produce national estimates on health care utilization. Section 3 provides information on the number of sampled health centers that were eligible to participate in the NAMCS HC Component and submitted data in 2023. Section 4 details the contents of the 2023 NAMCS HC Component public use data file and the edits used in the creation of the file.

Section 5 contains an explanation of the procedures used to accurately produce variance estimates.

NCHS presentation standards for proportions, counts, and rates, and their relation to NAMCS HC Component data, are discussed in Section 6, and the data analysis guidelines are provided in Section 7. Section 8 provides information on item missingness, and Section 9 provides a comparison of frequencies between the NAMCS HC Component public use and restricted use data files. Section 10 provides a list of preferred reporting items for complex sample survey analysis. Section 11 provides further information on the availability of NAMCS HC Component restricted use data files available in NCHS and Federal Research Data Centers. Appendix A provides unweighted frequencies for selected variables included on the public use data file.

Suggested Citation

Technical Documentation: National Center for Health Statistics. Division of Health Care Statistics. *2023 National Ambulatory Medical Care Survey Health Center (NAMCS HC) Component Public Use Data File Documentation*, August 2025. Hyattsville, Maryland.

Data File: National Center for Health Statistics. Division of Health Care Statistics. 2023 National Ambulatory Medical Care Survey Health Center (NAMCS HC) Component public use data file. 2025. Hyattsville, Maryland.

Contact Information

Data users can find the latest information about the NAMCS HC Component on our website, at: https://www.cdc.gov/nchs/ahcd/namcs_index.htm. If data users have queries about the public use data file, they may send their question through email to ambcare@cdc.gov, or call us at 301-458-4600. A response to data user inquiries is generally provided in 1-2 business days.

The National Center for Health Statistics has an ambulatory health care data listserv, where updates and information about the most recent ambulatory care data (including the NAMCS HC Component) are sent out. Details on how to subscribe to the NCHS Listserv for ambulatory health care data can be found at: https://www.cdc.gov/nchs/ahcd/ahcd_listserv.htm.

Contents

Section 1 About the National Ambulatory Medical Care Survey Health Center Component	6
Section 1.1 Background	6
Section 1.2 Data Sources	7
Section 1.3 Data Confidentiality.....	7
Section 2 Methodology	8
Section 2.1 Brief Overview	8
Section 2.2 Health Center Frame and Sample Design	8
Section 2.3 NAMCS HC Component Public Use Data File Sample Design.....	9
Section 2.4 Data Collection Procedures	10
Section 2.5 Weighting	10
Section 3 Sample Size, Eligibility, and Response Rate	12
Section 4 Data Processing	13
Section 4.1 Diagnosis Data.....	13
Section 4.2 Patient Age	14
Section 4.3 Patient Sex.....	14
Section 4.4 Patient Race and Hispanic Ethnicity	14
Section 4.5 Patient Marital Status.....	14
Section 4.6 Visit Month and Day	14
Section 5 Standard Errors and Variance Estimation	15
Section 5.1 Subpopulation Analysis – Subsetting Data	15
Section 6 Presentation Standards	16
Section 7 Data Analysis Guidance	17
Section 7.1 Visit weight	17
Section 7.2 Guidance on Weight Normalization	18
Section 7.2.1 Normalization Example	19
Section 7.3.1 Normalization Example Code	22
Section 7.3 SAS SUDAAN Survey Procedures	24
Section 7.3.1 NEST Statement Variables	24
Section 7.4 SAS Survey Procedures	25
Section 7.5 R Survey Procedures	25
Section 7.6 Stata Survey Procedures	26
Section 8 Survey Content	26

Section 8.1 Demographic Item Missingness Rate	27
Section 8.2 Diagnosis Item Missingness Rate	27
Section 9 Data Comparison	29
Section 9.1 Public Use Data Files and Restricted Use Data File.....	29
Section 10 Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA) Checklist for the 2023 NAMCS HC Component Public Use Data File	32
Section 11 Research Data Center	33
Section 12 References.....	34
Appendix A Unweighted frequencies for health center visits	35
Appendix B Health Center (HC) Facility Interview	37

Section 1 About the National Ambulatory Medical Care Survey Health Center Component

Section 1.1 Background

The National Ambulatory Medical Care Survey Health Center (NAMCS HC) Component is an annual survey that provides data on health care utilization at health centers in the United States. As a part of NAMCS, the National Center for Health Statistics (NCHS) began collecting data from health centers in 2006. A separate sample of health centers was drawn in 2012 for NAMCS. In 2021, NCHS redesigned the NAMCS HC Component to collect visit data from electronic health records (EHRs) from participating health centers for the entire calendar year.

The NAMCS HC Component collects data on health center visits including information on diagnoses and patient demographics. The survey aims to provide health trends and outcomes of the U.S. population's utilization of health centers in the following ways:

- Provide nationally representative, accurate, and reliable health care data for health centers in the United States.
- Answer key questions of interest to health care professionals, researchers, and policy makers about health care quality, use of resources, and disparities of services to population subgroups.
- Monitor national trends in health care topics for which health centers play an important role, such as mental health and substance use-related care, maternal and child health, and HIV-related care.
- Contribute to a stronger public health foundation that helps address current and future public health threats.

In 2022, the entire sample included 324 federally qualified health centers (FQHC) and FQHC look-alikes in the 50 U.S. states and the District of Columbia that used an EHR system. Out of the entire sample, 104 health centers were included in the primary sample and 220 health centers made up the reserve sample. Ultimately, 315 health centers were contacted and 95 health centers agreed to participate and provided visit data from their EHRs. Out of the 95 responding health centers, 63 health centers continued participation from 2022, 27 health centers were newly recruited in 2023, and five other health centers that participated in 2023 were first contacted in prior years, but did not provide EHR data in 2022. For more detailed information regarding the sample frame, see Section 2.2.

Overall, 9,012,885 health centers visits were collected from the 95 responding health centers. Of these, 450,645 health center visits were selected to create the 2023 NAMCS HC Component public use data file.

Section 1.2 Data Sources

The NAMCS HC Component receives data from EHR systems. Participating health centers submit EHR data, which contain an unlimited number of medical diagnosis and procedure codes, laboratory and medication data, and unstructured clinical notes. However, the public use data file will only include diagnosis variables and demographic information. The NAMCS HC Component accepts EHR data in the format of HL7 CDA® R2 Implementation Guide: National Health Care Surveys Release 1, DSTU Release 1.2 – US Realm (http://www.hl7.org/implement/standards/product_brief.cfm?product_id=385).

However, some EHR vendors are not able to format their data in the HL7 CDA format as specified in the National Health Care Surveys Implementation Guide. Alternatively, these centers were able to submit their EHR data as custom extracts, which contained many (but not all) data elements extracted via the above implementation guide.

Section 1.3 Data Confidentiality

NCHS and its designated agents take the security and confidentiality of NAMCS HC Component public use data file very seriously. Strict laws have been implemented to establish minimum Federal standards for safeguarding the privacy of individually identifiable health information. Assurance of confidentiality is provided to all health centers according to Section 308(d) of the Public Health Services Act [42 United States Code 242m (d)]. Strict procedures according to Section 3572 of the Confidential Information Protection and Statistical Efficiency Act (44 U.S.C. 3561-3583) are utilized to prevent disclosure of personal identifiable information in NAMCS HC Component data. All information which could identify a participating health center is confidential and seen only by persons associated with NAMCS HC Component, and is not disclosed or released to others for any other purpose. Prior to the release of public use data file, NCHS conducts extensive disclosure risk analysis to minimize the chance of inadvertent disclosure. As a result, selected characteristics and/or data elements may have been omitted or masked on the public use data file to minimize the potential risk of disclosure. Masking was performed in such a way to cause minimal impact on the data. See Section 4: Data Processing for more information on which data elements in the public use data file were impacted.

The protocol for NAMCS HC Component has been approved by the NCHS Research Ethics Review Board since the survey's establishment (2006).

Section 2 Methodology

Section 2.1 Brief Overview

The 2023 NAMCS HC Component used a national probability sample of health centers to collect data on visits to develop the public use data file. The 2023 NAMCS HC Component public use data file sample was designed to allow for nationally representative estimates of visits at health centers in the United States.

Section 2.2 Health Center Frame and Sample Design

The 2023 NAMCS HC Component identified a targeted universe of FQHCs and FQHC look-alikes in the 50 U.S. states and the District of Columbia that provide direct ambulatory care and use an EHR system at one or more delivery sites. Health centers that were fully or partially funded by the Health Resources and Services Administration (HRSA) were considered for inclusion. Health centers were deemed ineligible if they:

- Did not have an EHR system
- Did not provide healthcare services to the general U.S. population, or only provided care to special institutionalized populations such in prisons, nursing homes, homeless shelters, etc.
- Only provided dental services
- Were located on a military installation or outside of the 50 U.S. states and the District of Columbia

To create the sampling frame and draw the sample, NCHS worked with the HRSA to use a nationally representative database that contains a list of all health centers in the United States. The database contained 1,487 health centers for the 2023 NAMCS HC Component. To create the sampling frame from this database, ineligible health centers were removed. This included 48 health centers that did not meet the inclusion criteria described above and 322 health centers that were included in the 2021 sample. This process yielded a sampling frame of 1,117 eligible health centers.

In 2021, a stratified random sample of 50 FQHCs and FQHC look-alikes was drawn as the primary sample, along with a reserve sample of 100 health centers. The 2022 NAMCS HC Component sample was expanded to initially add 60 respondent health centers to the 50 respondent health centers from the

2021 sample, resulting in 110 FQHCs and FQHC look-alikes making up the 2022 NAMCS HC Component sample. However, 54 health centers were ultimately fielded due to budget constraints. Due to this, six randomly selected health centers were removed from the sample in four strata. In 2022, an additional 120 additional health centers were selected for the reserve sample (Williams et al., 2023). In 2023, 50 health centers were selected for the primary sample and 100 health centers for the reserve sample.

Ultimately 315 health centers were contacted to participate in the 2023 NAMCS HC Component, which includes 95 respondents and 220 eligible non-respondents. The 95 participating health centers include 26 health centers from the 2021 sample, 38 health centers from the 2022 sample, and 31 health centers from the 2023 sample. Weighting was conducted to produce health center-level and visit-level estimates. Data were collected for 100% of visits from the sampled health centers via EHR submission.

Section 2.3 NAMCS HC Component Public Use Data File Sample Design

While the NAMCS HC Component restricted use data file includes every health center (HC_j) visit record submitted to NAMCS HC Component for the survey year, the 2023 NAMCS HC Component public use data file consists of a 5% sample of NAMCS HC Component visit data. This 5% sample of NAMCS HC Component records was selected for the public use data file instead of the full listing of records to decrease disclosure risk and increase efficiency for data users when conducting statistical analyses.

In 2023, the NAMCS HC Component collected 9,012,885 visit records. Stratified systematic sampling was used to select the public use data file sample of health center visits. A targeted number of records was determined by taking 5% of the total health center visit records ($n=450,645$). The sampling interval was the inverse of the percent of submitted EHRs targeted for inclusion in the subsample. The sampling interval used to select the public use data file records in the 2023 NAMCS HC Component was $1/0.05$, or 20. Within each estimation stratum, participating health centers were randomly ordered. Within each health center, visits were then sorted by the following variables:

Visit Week → Day of Week

Once sorted, visits were serially numbered in each estimation stratum. Next from the ordered array of HC_j records, visits were selected for the public use data file sample if the assigned “array sequential” numbers were the nearest integer greater than or equal to:

$$R_r + \text{Int}(EHR)_r \times k$$

Where:

R_r = random number between 0 and $Int(EHR)_r$

$k = 0, 1, 2, 3, \dots$

$Int(EHR)_r$ = sampling interval

Section 2.4 Data Collection Procedures

In 2023, health centers submitted EHR data via two sources, either directly from the health centers' EHR system or as a custom extract, as mentioned above in Section 1.2. Once data were collected, several steps were required for data processing. Specifications for checking, configuring, and transmitting the data files were developed by NCHS. Once NCHS received the data files they were processed to harmonize data from the two data sources. All records from participating health centers' EHRs were brought into the restricted database, and those records were then collapsed so that a given patient could only have one record (called a visit in the PUF) per day at a given health center.

Section 2.5 Weighting

Weighting was conducted to produce health center-level and visit-level estimates, and to account for sampling probabilities and nonresponse. Only visit-level weights are included in the public use data file, and users are only able to produce visit-level estimates with this file.

Health center-level data were collected via self-completed forms from participating health centers (Appendix B). All 2023 health center visits were collected from the sampled health centers via electronic files of their EHR system. Participating health centers submitted data for all visits that occurred during the 2023 calendar year. While the 2023 NAMCS HC Component restricted use data file includes all (100%) of the visit records sent, the public use data file includes a 5% sample of those records, as described in Section 2.3.

All health center visit data collected for 2023 were used to develop weights. To produce visit-level weights, health center-level weights were first developed and smoothed. The visit-level weights were then developed for the restricted use file that includes all visits from participating health centers. These visit weights were formulated as the final health center weight multiplied first by the health center's actual annual number of visits made for medical care followed by a partial non-response adjustment factor. Visit weights for all visits were then smoothed before they were finalized. Because the public use data file only contains a 5% sample of all visits submitted in the 2023 NAMCS HC Component, visit

weights for visits included on the public use data file were adjusted accordingly. This ensures that weighted estimates from the restricted use file and the public use data file sum to approximately the same total number of weighted visits at health centers in the survey year.

Variance estimation procedures for weighted estimates are described further in Section 5 with coding examples in Section 7, and comparisons of weighted estimates between the restricted and public use data files in Section 9.

Section 3 Sample Size, Eligibility, and Response Rate

All 315 health centers that were contacted for participation were eligible to participate in the survey. Ultimately, 95 health centers participated in the 2023 NAMCS HC Component yielding a response rate of 30.2%. In 2023 of the health centers that were newly recruited, 64 health centers were first contacted for participation, and 27 of these health centers agreed to participate and submitted EHR data (42.2%). Of the 64 health centers that previously participated in 2022, 63 were retained and continued participation in 2023 (98.4%). Five other health centers that participated in 2023 were first contacted in prior survey years, but did not provide EHR data in 2022. A health center was considered a full respondent if they provided data for at least six months of the survey year. Of the 95 participating health centers that were included in the 2023 NAMCS HC restricted use data file, all provided at least six months of data. Therefore, all health centers were selected to create the public use data file. From the 95 health centers, 5% of all records were selected for the public use data file. Overall, 450,645 health center visits were selected. Table 3.1 presents the response rates and retention rates for the 2022-2023 NAMCS HC Component. Table 3.2 presents the number of health centers and visits for the 2023 NAMCS HC Component.

Table 3.1 Response rates and retention rates, NAMCS HC Component, 2022-2023

	Unweighted overall response rate	Year-specific response rate	Retention Rate
2022	25.1	26.1	89.7
2023	30.2	42.2	98.4

Note: The unweighted overall and year-specific response rates were calculated using American Association for Public Opinion Research (AAPOR) Response Rate 4 formula. The unweighted overall response rate is a calculation of the eligible and partial respondents divided by the eligible respondents, partial respondents, and eligible non-responding health centers. The year-specific response rate is a calculation of the eligible and partial respondents that were newly recruited in a specific year divided by the eligible respondents, partial respondents, and eligible non-responding health centers for that specific year. The retention rate is a calculation of the number of health centers who originally participated in the NAMCS HC Component in year(s) prior and continued to participate in 2023 divided by the number of health centers who originally participated in the NAMCS HC Component in year(s) prior.

Table 3.2 Number of health centers and visits, NAMCS HC Component, 2023

	Health Centers	Visits
Restricted Use Data File	95	9,012,885
Public Use Data File	95	450,645

Section 4 Data Processing

The data included in the public use data file underwent additional processing to prepare them for release. Suppression rules such as masking were applied for some records to protect patient confidentiality. Other items were either top-coded or bottom-coded in accordance with NCHS confidentiality requirements; this is noted for specific data items outlined in this section. Imputation was not conducted for data elements with missing values prior to creation of the 2023 NAMCS HC Component public use data file.

Section 4.1 Diagnosis Data

In the 2023 NAMCS HC Component, diagnosis data from participating health centers were submitted in three different diagnosis coding systems including: *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)*; *International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM)*; and *SNOMED Clinical Terms (SNOMED CT)*. In the creation of a harmonized and integrated database, the ICD-9-CM and SNOMED CT diagnosis codes were translated to ICD-10-CM, where applicable. Translation from ICD-9-CM and SNOMED CT to ICD-10-CM was the only modification to the diagnosis codes. On the public use data file, medical diagnosis codes were limited to ICD-10-CM diagnosis codes.

An ICD-10-CM code can have a maximum of 7 characters and is organized by chapters from A to Z. For the 2023 NAMCS HC Component public use data file, ICD-10-CM codes have been truncated to four characters to minimize disclosure risks. While the codes have been truncated, the diagnosis codes are never updated or revised to a different code that would result in a change to the original diagnosis for a visit. To maintain integrity of the data, any codes that appear to be invalid are kept as is.

Duplicate 4-character ICD-10-CM codes were removed for each unique visit on the public use data file. Although visits collected from health center EHR systems could have had an unlimited number of diagnosis records, diagnosis codes were limited to 30 unique codes per visit (variables DX1 through DX30) in the public use data file, which captured 96.6% of diagnoses recorded at visits included on the public use data file. Rarity of diagnoses was assessed and those deemed rare were truncated to two characters.

At least one diagnosis code is listed in 67.7% of all visits. Four health centers did not provide any condition codes that could be translated to ICD-10-CM, therefore do not have any visits that include at

least one condition code in DX1-DX30. Of the 91 health centers that provide any codes that translated to ICD-10-CM, 69.3% of their visit have at least one diagnosis code in the public use data file.

Section 4.2 Patient Age

Patient age is present for all visits in the 2023 NAMCS HC Component public use data file. Visits were top coded to the 99.5th percentile of age, thus visits by patients ages 88 and older were top coded to 88 years.

Section 4.3 Patient Sex

Patient sex is missing in 0.7% of records on the 2023 NAMSC HC Component public use data file.

Section 4.4 Patient Race and Hispanic Ethnicity

Patient race is missing from 28.0% of records on the 2023 NAMCS HC Component public use data file. Eighteen health centers are missing patient race for all visit records. Excluding the 18 health centers with complete missingness, 18.2 % of visits are missing patient race.

Patient ethnicity is missing from 22.3% of records on the 2023 NAMCS HC Component public use data file. Eighteen health centers are missing patient ethnicity for all visit records. Excluding the 18 health centers with complete missingness, 11.7% of visits are missing patient ethnicity.

Section 4.5 Patient Marital Status

Marital status of patients is included in the public use data file but is missing from 26.2% of records overall. Fourteen health centers are missing marital status from all visit records. For the remaining 81 health centers, marital status is missing from 18.9% of visits.

Section 4.6 Visit Month and Day

Exact dates are not provided on the NAMCS HC Component public use data file. Instead, only the month and day of the week of health center visits are provided.

Section 5 Standard Errors and Variance Estimation

Standard error is primarily a measure of the sampling variability that occurs by chance because only a sample of health centers are in NAMCS HC Component, rather than the entire universe of health centers. Standard errors and other measures of sampling variability are best determined by using a statistical software package that takes into account the sample designs of surveys to produce such measures.

See Section 7 for further guidance on how to apply weights and calculate standard errors to generate national estimates.

Section 5.1 Subpopulation Analysis – Subsetting Data

For data users who may have a subpopulation of interest, such as a particular age group or sex, a domain analysis must be performed, also known as a subgroup or subpopulation analysis.

For some variance estimation methods, the entire set of data containing the appropriate weights for a particular survey year must be used to obtain the correct variance estimates. Therefore, it is not recommended to drop observations from the dataset when subsetting data, as it may affect variance estimation. Instead, the estimation procedure must indicate which records are in the subgroup of interest. For example, when examining female patients aged 35 and over, the entire dataset of examined individuals (both male and female patients of all reported ages) must be read into the statistical software program.

The STAT and DOMAIN statements in the SAS survey procedure, SUBPOP in SAS callable SUDAAN, or comparable statements in other programs (SUBSET in R; subpop or over in Stata) must be used to indicate the subgroup of interest (i.e., females aged 35 and over).

Depending on the specifications of a data user's statistical software of choice, an indicator variable created by the data user prior to the procedure may facilitate the identification of the subgroup in the procedure statements.

Section 6 Presentation Standards

Data users should be aware of the reliability of survey estimates, particularly smaller estimates. NCHS has published standards for the assessment of reliability and presentation of proportions (or percentages) (https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf) and for the presentation of rates and counts (https://www.cdc.gov/nchs/data/series/sr_02/sr02-200.pdf). For presentation or publication of count estimates using data from the NAMCS HC Component, we recommend visit estimates be rounded to the nearest thousand.

These presentation standards apply to products published by NCHS. If, according to the presentation standards, an estimate is not reliable, data users should examine the confidence interval carefully before using the estimate.

Section 7 Data Analysis Guidance

The following section provides an overview on how data users can derive visit estimates and compute variances to produce standard errors, using statistical software tools such as SAS, R, and Stata. For the NAMCS HC Component public use data file, SAS-callable SUDAAN software procedures are used for survey analysis, however, SAS/STAT software procedures beginning with SURVEY for survey analysis may also be used. R relies on the “survey” package to conduct survey data analysis whereas Stata, uses the “svy” command. SAS/SUDAAN, R and Stata users can use these procedures to conduct statistical analysis on data from the 2023 NAMCS HC Component public use data file. Additionally, this section provides guidance on normalizing visit weights to account for complete missingness for analytic variables of interest. The guidance provides data users a framework to implement normalizing weights for data analysis. Data users should always investigate if there are any variables of interest that have complete missingness at health centers in the 2023 NAMCS HC Component public use data file.

Section 7.1 Visit weight

The visit weight is a critical component in the process of producing estimates from sample data and its use should be clearly understood by all data users. The statistics contained on the public use data file reflect only a sample of visits; a 5% sample of the NAMCS HC Component data collected from participating health centers, not a complete count of all visits that occurred in the United States. Each health center’s visit record in the public use data file represents one patient visit in the sample of 450,645 visits. To obtain national estimates from the 5% sample, each record is assigned an inflation factor called the “visit weight” (variable VISWT in the public use data file).

By aggregating the “visit weights” assigned to the VISWT variable on the 450,645 health center visits for 2023, the data user can obtain the estimated total of 119,939,235 health center visits (standard error of 16,469,993 health center visits) made in the United States in 2023.

Note that estimates of health center visits produced from the 2023 NAMCS HC Component public use data file may differ somewhat from those estimates produced from the 2023 NAMCS HC Component restricted use data file. This is because of adjustments required for the public use data files, as part of the disclosure risk mitigation process. Certain variables were masked on some records for confidentiality purposes. Other variables were top and/or bottom coded in accordance with NCHS confidentiality requirements.

The table in Section 9 compares aggregate unweighted and weighted data for selected variables between the 2023 NAMCS HC Component public use data file and restricted use data file.

Section 7.2 Guidance on Weight Normalization

Some health centers did not provide certain data elements for any of their visits in the 2023 data year. In certain situations, some health centers needed to produce custom extracts of their records to conform with the format needed for processing as specified in the HL7 CDA Implementation Guide. Therefore, not all data elements were required of health centers providing custom extracts. In other situations, even for health centers providing data via the IG, certain variables were incomplete for all visits at specific health centers.

Regardless of the reason for missingness, data users must identify health centers that have complete missingness for specific analytic variable(s) of interest, and exclude those health centers' visits from analysis. Additionally, if certain health centers' visits must be excluded, users must normalize the weight variable (VISWT) so that the sum of weights of visits in the analysis is equal to the sum of weights of all visits in the 2023 NAMCS HC Component public use data file.

Steps for a complete case analysis:

1. Identify health centers to be included in your analysis:
 - a. Identify variable(s) required for your analysis
 - b. Identify health centers that are missing values at ALL visits for at least one variable of interest from Step 1a
 - c. Subset all visits from health centers identified with complete missingness for at least one variable of interest, as identified in Step 1b above.

NOTE: This process does not eliminate all missingness, rather it eliminates complete missingness of a specific variable for a specific health center. Health centers that are included may still have some visits with missing information for the variables of interest, but this process removes visits at health centers that did not provide any information for variables of interest.

2. Normalize weights if only a subset of health centers' visits is included:
 - a. Calculate the sum of weights for all visits in the public use data file. In 2023, the sum of weights (VISWT) is 119,939,235.
 - b. Calculate the sum of weights for visits at health centers to be included in your analysis.

- c. Calculate the normalization factor [X] by dividing the sum of weights for all visits in the survey (from step 2a) by the sum of weights for visits in your analysis (from Step 2b), and the value of X from this calculation is the factor you will use to normalize your weights.
 - i. $X = [\text{sum of all visit weights}] / [\text{sum of visit weights in your analysis}]$
 1. NOTE: X will always be greater than 1.
- d. Create a new weight variable for visits in your analysis by multiplying the original weight variable by your normalization factor (X).
 - i. $\text{NEW_WT} = \text{VISWT} * X$
- e. Use NEW_WT for your analysis in place of VISWT.

NOTE: If you add or subtract variables from your analysis, or you develop a new research question and analysis, you must conduct these steps again to ensure that you: 1) capture visits from health centers providing data on your variables of interest, and 2) normalize those visits' weights accordingly.

Table 7.1 Variables that contain health centers with complete missingness in the 2023 NAMCS HC public use data file

Variable Name	Variable Description	HCID_5 to exclude
DX1-DX30	Diagnoses 1-30	15, 71, 83, 90
ETHNICITY	Patient Hispanic ethnicity	18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 82, 85, 88
MARITAL	Marital status	18, 20, 21, 24, 29, 42, 44, 49, 51, 58, 63, 82, 85, 88
RACE	Patient race	18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 82, 85, 88
RACERETH	Combined race and ethnicity variable	18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 82, 85, 88

Section 7.2.1 Normalization Example

The example below will showcase the differences in estimates when normalizing the 2023 NAMCS HC public use data file for visits with a mental health disorder and race as opposed to not normalizing. This example will provide context on normalizing weights when assessing complete missingness for two variables on the public use data file (DX1 and RACE).

Before following the steps for a complete case analysis, it is helpful to assess the unweighted and weighted number of visits for all 95 health centers in the public use data file, as shown in Table 7.2. There are 450,645 visits in the public use data file representing 119,939,235 health center visits.

Table 7.2 Weighted and unweighted number of visits in the 2023 NAMCS HC Component public use data file

Visits at all health centers (N=95)	
Unweighted	450,645
Weighted	119,939,235

In this example, assume the user wants to assess visits with a first-listed diagnosis (DX1) of a mental health disorder, stratified by race (RACE) using the 2023 NAMCS HC Component PUF. For the purposes of this example, a mental health disorder was classified as any ICD-10-CM code in the Mental, Behavioral and Neurodevelopmental disorders chapter (F01-F99). Please note that in this public use data file, when DX1 is missing, all DX1-DX30 variables will be missing, so whether assessing first-listed or any-listed diagnosis, we only need to assess complete missingness for DX1.

First, the user must identify all health centers that have complete missingness in *either* the race (RACE) or first-listed diagnosis (DX1) variables (or both) from Table 7.1 above. In 2023, 22 health centers have complete missingness in the DX1 or RACE variables. Health centers 15, 71, 83, 90 are missing DX1 at all visits. Health centers 18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 82, 85, 88 are missing RACE at all visits. Therefore, 73 health centers make up the subset of data to analyze first-listed mental health diagnoses by race. The normalization factor X should be calculated by dividing the sum of all visit weights (119,939,235) by the sum of visit weights from the 73 health centers included in this example (101,694,007). The normalization factor is $(119,939,235/101,694,007)$ or approximately 1.2. The normalization factor is used to create a new weight variable, which for this example is calculated as $NEW_WT = VISWT * (1.2)$. After calculating the normalization factor and creating a new weight variable, the data user should apply the new visit weight variable to the subset of visits at the 73 health centers included in this example. The total sum of weights in the analytic subset (sum of NEW_WT at HC visits to be included) should be equal to the total sum of weights for all visits at all health centers in the NAMCS HC public use data file as shown in Table 7.2.

At the 73 health centers identified for inclusion in this example, we identified visits with a first-listed mental health ICD-10-CM diagnosis and race information. We then produced unweighted and weighted estimates (using the normalized NEW_WT variable) of visits with a first-listed mental health diagnoses

at health centers in 2023. These estimates are detailed in Table 7.3 for users to replicate. Please note, normalization of weights at the subset of visits to be included only impacts the weighted numerator and weighted denominator estimates; the unweighted counts and weighted percentage will not change in the same subset of visits due to weight normalization.

Table 7.3 Visits with a first-listed mental health diagnosis, with race and diagnosis information, in the 2023 NAMCS HC public use data file

	Overall	Non-Normalized subset	Normalized subset
Analysis	Overall Data File	Subset without Normalization implemented	Subset with Normalization correctly implemented
Number of health centers	95	73	73
Unweighted numerator	38,958	32,760	32,760
Unweighted denominator	450,645	386,014	386,014
Weight used	VISWT	VISWT	NEW_WT ¹
Weighted numerator	9,726,203	8,125,301	9,583,086
Weighted denominator	119,939,235	101,694,007	119,939,235
Weighted Percent (SE)	8.1 (0.9)	8.0 (1.0)	8.0 (1.0)

¹ As described in Section 7.2.1, $NEW_WT = VISWT * 1.2$, where 1.2 is the calculated normalization factor.

In the first column of Table 7.3, the data is neither subset nor using a normalized visit weight. The weighted numerator underestimates the weighted number of visits with a first-listed mental health diagnosis and race, which also results in an underestimated weighted percent. In the second column, the data is subset to exclude health centers with complete missingness but does not use the normalized visit weight. This further underestimates the weighted number of visits with a first-listed mental health diagnosis. Additionally, because of the use of the subset of health centers and a non-normalized visit weight in the second column, the weighted denominator does not add up to the total number of visits in the public use data file. The last column displays the correct analysis using the subset of health centers and the normalized weight variable.

Using a subset of health centers and normalizing their weights produces a higher weighted numerator than using all health centers and the non-normalized weight or using the subset of health centers and the non-normalized weight. In the overall analysis in Table 7.3, visits at health centers with complete missingness for diagnosis data are automatically considered to be non-mental health visits despite not having enough information to discern whether there was a mental health diagnosis. Consequently, the

overall weighted numerator is an undercount of visits with a first-listed mental health diagnosis at health centers in the United States.

In short, normalizing weights may produce different estimates when analyzing the 2023 NAMCS HC Component public use data file depending on the number of health centers that are included in the analysis. Without excluding health centers with complete missingness and subsequently normalizing visit weights, data users will underreport counts and rates for their analysis of interest. Data users should consider the full scope of their research question to make decisions on the subset of health centers to include, and how normalizing visit weights will impact the calculation of estimates. Data users should reference Table 7.1 to ensure the correct health centers are excluded in their analysis when normalizing weights in a complete case analysis.

Section 7.3.1 Normalization Example Code

For further assistance in implementing normalization on the 2023 NAMCS HC Component public use data file, the following SAS code replicates the normalization example described in Section 7.2.1.

*STEP 1;

*Identify the variables of interest for your analysis;

 *Research Question: First listed diagnoses of mental health by age and race;

 *Variables needed: DX1, RACE;

*In this example you will need to subset the data where DX1 is missing or RACE is missing according to Table 7.1;

 *DX1 is missing where HCID_S in (15,71,83,90);

 *RACE is missing where HCID_S in (18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 82, 85, 88);

*STEP 2;

*Calculate two sums:

 1. the sum of weights at all HCs in the original datafile and

 2. the sum of weights at HCs to be included in your analysis;

*1. Overall sum of weights;

proc sql;

 create table sum_total as

 select sum(viswt) as sum_total

 from /*[full datafile]*/;

quit;

proc print data=sum_total;

run;

*2. Subset sum of weights;

```

proc sql;
    create table sum_subset as
    select sum(viswt) as sum_subset
    from /*[full datafile]*/
    where HCID_S not in (15, 18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 71, 82, 83, 85,
    88, 90);
quit;
proc print data=sum_subset;
run;

*STEP 3;
*Create two new variables for your analysis:
    1. a normalized weight, using sum_total and sum_subset calculated in step 4 and
    2. an inclusion indicator where the record is at a PSU identified in 'all_three' from STEP 3 above;
data /*new_datafile*/;
    set /*[full datafile]*/;
        new_wt=viswt*(/*[value of sum_total]/[value of sum_subset]*/);
        if HCID_S not in (15, 18, 19, 20, 21, 24, 29, 33, 37, 42, 44, 49, 51, 58, 63, 68, 71, 82, 83,
        85, 88, 90)
            then include=1;
            else include=2;
        if "F01"<=substr(DX1, 1, 3)<="F99"
            then mntlhlth=1;
            else mntlhlth=0;
run;

*STEP 4;
*Use these two variables (new_wt and include) in all procedures used to produce weighted output for
this analysis;
*Note: this step shows a SUDAAN procedure for setting up a weighted analysis, but an example of a SAS
procedure is provided below in Section 7.4;
*First, sort the data by STRATUM and HCID_S;
proc sort data=/*[new datafile]*/;
    by STRATUM_S HCID_S;
run;

*Second, set up your SUDAAN statement as follows (showing a crosstab procedure);
proc crosstab data=[new_datafile] filetype=sas design=wr atlevel1=1 atlevel2=2;
    nest STRATUM_S HCID_S / MISSUNIT;
    weight new_wt;
    subpopn include=1;
    class mntlhlth; *include analytic indicators/variables to cross;
    tables mntlhlth; *cross your class variables in the desired order;
    output nsum wsum sewgt totper setot atlev1 atlev2 / filename = /*[output dataset]*/ replace
tablecell=default;
run;

```

Section 7.3 SAS SUDAAN Survey Procedures

The program below demonstrates how to set up your design and weight variables to produce weighted estimates using the 2023 NAMCS HC Component public use data file:

```
PROC (procedure) DATA=(input data set) FILETYPE=SAS DESIGN=WR ATLEVEL1=1 ATLEVEL2=2;
NEST STRATUM_S HCID_S / MISSUNIT;
*SUBPOPN (variable1) = (value); *Only use subpopn statement if needed;
WEIGHT VISWT; *or replace with your normalized weight if required for your analysis;
CLASS (variable2);
TABLES (variable2);
OUTPUT nsum wsum sewgt totper setot atlev1 atlev2 / FILENAME=[output dataset] REPLACE
TABLECELL=DEFAULT;
RUN;
```

In the above example, replace the parentheses with the information named in the parentheses. When health centers are missing a data element for all visits, those health centers' visits should be excluded from your analysis. If a subset of health centers' visits must be excluded due to complete missingness, replace VISWT with normalized version of the weight, and add a SUBPOPN statement to correctly subset to health centers' visits of interest. Refer to Section 7.2 above, for more guidance on weight normalization to account for complete missingness.

When using SAS-callable SUDAAN software, sort the input data set in the order specified in the NEST statement, in this case by sampling strata (STRATUM_S) followed by health center identifier (HCID_S) within STRATUM_S. If software other than SUDAAN is used to approximate the variances, other statements will be required by that software. The variance variables required by that software are the same as those include in the above example, which are further explained below in Section 7.3.1.

Section 7.3.1 NEST Statement Variables

The SUDAAN NEST statement for variances at the visit-level is:

```
NEST STRATUM_S HCID_S / MISSUNIT;
```

Where:

STRATUM_S is the scrambled value of the original sampling stratum from which the health center was selected.

HCID_S is the scrambled identifier for the health center.

Section 7.4 SAS Survey Procedures

The program below demonstrates how to calculate variance estimates using SAS SURVEYFREQ and SURVEYMEANS procedures:

For categorical variables:

```
PROC SURVEYFREQ DATA = (input data set);  
TABLE VAR1; *Replace "VAR1" with the categorical variable of interest.  
CLUSTER HCID_S;  
STRATA STRATUM_S;  
WEIGHT VISWT; *or replace with your normalized weight if required for your analysis;  
ODS OUTPUT ONEWAY=(name of output);  
RUN;
```

For continuous variables:

```
PROC SURVEYMEANS DATA = (input data set);  
VAR VAR1; *Replace "VAR1" with the continuous variable of interest.  
CLUSTER HCID_S;  
STRATA STRATUM_S;  
WEIGHT VISWT;  
ODS OUTPUT STATISTICS=(name of output);  
RUN;
```

In the above example, replace the parentheses with the information named in the parentheses. When health centers are missing a data element for all visits, those health centers' visits should be excluded from your analysis. If a subset of health centers' visits must be excluded due to complete missingness, replace VISWT with normalized version of the weight, and add a DOMAIN statement to correctly subset to health centers' visits of interest. Refer to Section 7.2 above, for more guidance on weight normalization to account for complete missingness.

Section 7.5 R Survey Procedures

The R package "survey" can be utilized for complex survey analysis (<https://cran.r-project.org/web/packages/survey/index.html>). The R programs below demonstrate how to install the survey package, produce visit level weighted estimates, and calculate variance estimates.

```
install.packages("survey")  
library(survey)  
install.packages("tidyverse")  
library(tidyverse)
```

```
#Using the “survey” package:  
{variable name} <- svydesign(  
  ids = ~ HCID_S,  
  strata = ~ STRATUM_S,  
  weights = ~ VISWT,  
  data={input data frame})
```

Note: Replace curly brackets {} with the information named in the parenthesis

Section 7.6 Stata Survey Procedures

The Stata programs below demonstrate how to use visit weights and calculate variance estimates with the svyset command (<https://www.stata.com/manuals/svysvyset.pdf>)

For categorical variables:

```
/*Set survey design*/  
svyset HCID_S [pweight = VISWT], strata(STRATUM_S)  
  
/*Specify one-way tables, change “VAR1” to categorical variable of interest*/  
svy: tab VAR1, count se  
svy: tab VAR1, percent
```

For continuous variables:

```
/*Set survey design*/  
svyset HCID_S, [pweight= VISWT], strata(STRATUM_S)  
  
/*Specify one-way tables, change “VAR1” to continuous variable of interest*/  
svy: mean VAR1
```

Section 8 Survey Content

For the 2023 NAMCS HC Component public use data file, 75 variables were included; 60 (77.9%) variables include information on medical diagnoses, 8 (10.4%) variables include patient demographic information, 2 (2.6%) data items include visit information, and 7 (9.1%) variables include weights or other survey-related information.

Please refer to the 2023 NAMCS HC Component public use data file codebook for detailed information on the variables, including variable names, variable type, variable descriptions, and variable values. Additionally, unweighted frequencies for selected variables on the public use data file are available in Appendix A.

Section 8.1 Demographic Item Missingness Rate

In the 2023 NAMCS HC Component public use data file, four (5.2%) demographic variables had an unweighted missingness rate that was greater than 5% including RACE, ETHNICITY, RACERETH and MARITAL.

The variables in the table below had an unweighted item missingness percentage greater than 5%. As explained in Section 7.2, some health centers contained complete missingness in certain variables. In Table 8.1, two denominators are presented to demonstrate missingness. First, is the percent missing among all visits in all health centers (N=95) in the public use datafile. The second denominator is the percent missing among all visits in all health centers that do not have complete missingness in the diagnosis variable (N=81 or N=77).

Table 8.1 Percent missing (unweighted) for demographic variables in the NAMCS HC Component public use data file with a missingness greater than 5%

Variable Name	Variable Description	% Missing (all visits)	% Missing ¹
RACERETH	Patient Race and Ethnicity	19.2	8.2 ²
ETHNICITY	Patient Hispanic Ethnicity	22.3	11.7 ²
RACE	Patient Race	28.0	18.2 ²
MARITAL	Marital Status	26.2	18.9 ³

¹Denominators vary as percentages exclude health centers with complete missingness.

²N=77 health centers.

³N=81 health centers.

Section 8.2 Diagnosis Item Missingness Rate

In the 2023 NAMCS HC Component public use data file, 60 diagnosis variables (77.9%) had an unweighted missingness rate that was greater than 5%. It is expected that most of the diagnosis variables after the first-listed diagnosis variable will have a high missingness percentage as not all visits are expected to have multiple diagnoses.

The variables in the table below had an unweighted item missingness percentage greater than 5%. As explained in Section 7.2, some health centers contained complete missingness in certain variables. In Table 8.2, two denominators are presented to demonstrate missingness. First, is the percent missing among all visits in all health centers (N=95) in the public use data file. The second denominator is the percent missing among all visits in all health centers that do not have complete missingness in the diagnosis variable (N=91).

Table 8.2 Percent missing (unweighted) for diagnoses variables in the NAMCS HC Component public use data file with a missingness greater than 5%

Variable Name	Variable Description	% Missing (All visits)	% Missing ¹
DX1	Diagnosis #1 (ICD-10-CM), diagnosis code	32.4	30.7
DX2	Diagnosis #2 (ICD-10-CM), diagnosis code	51.9	50.7
DX3	Diagnosis #3 (ICD-10-CM), diagnosis code	63.2	62.3
DX4	Diagnosis #4 (ICD-10-CM), diagnosis code	71.2	70.6
DX5	Diagnosis #5 (ICD-10-CM), diagnosis code	77.3	76.8
DX6	Diagnosis #6 (ICD-10-CM), diagnosis code	82.3	81.9
DX7	Diagnosis #7 (ICD-10-CM), diagnosis code	85.7	85.4
DX8	Diagnosis #8 (ICD-10-CM), diagnosis code	88.2	87.9
DX9	Diagnosis #9 (ICD-10-CM), diagnosis code	90.1	89.8
DX10	Diagnosis #10 (ICD-10-CM), diagnosis code	91.6	91.4
DX11	Diagnosis #11 (ICD-10-CM), diagnosis code	92.7	92.6
DX12	Diagnosis #12 (ICD-10-CM), diagnosis code	93.7	93.5
DX13	Diagnosis #13 (ICD-10-CM), diagnosis code	94.4	94.3
DX14	Diagnosis #14 (ICD-10-CM), diagnosis code	95.1	94.9
DX15	Diagnosis #15 (ICD-10-CM), diagnosis code	95.6	95.5
DX16	Diagnosis #16 (ICD-10-CM), diagnosis code	96.1	96.0
DX17	Diagnosis #17 (ICD-10-CM), diagnosis code	96.5	96.4
DX18	Diagnosis #18 (ICD-10-CM), diagnosis code	96.8	96.7
DX19	Diagnosis #19 (ICD-10-CM), diagnosis code	97.1	97.1
DX20	Diagnosis #20 (ICD-10-CM), diagnosis code	97.4	97.3
DX21	Diagnosis #21 (ICD-10-CM), diagnosis code	97.6	97.6
DX22	Diagnosis #22 (ICD-10-CM), diagnosis code	97.9	97.8
DX23	Diagnosis #23 (ICD-10-CM), diagnosis code	98.1	98.0
DX24	Diagnosis #24 (ICD-10-CM), diagnosis code	98.2	98.2
DX25	Diagnosis #25 (ICD-10-CM), diagnosis code	98.4	98.3
DX26	Diagnosis #26 (ICD-10-CM), diagnosis code	98.5	98.5
DX27	Diagnosis #27 (ICD-10-CM), diagnosis code	98.6	98.6
DX28	Diagnosis #28 (ICD-10-CM), diagnosis code	98.8	98.7
DX29	Diagnosis #29 (ICD-10-CM), diagnosis code	98.9	98.8
DX30	Diagnosis #30 (ICD-10-CM), diagnosis code	99.0	98.9
DX_TYPE1	Diagnosis Type #1. Corresponds to Diagnosis #1	55.5	54.4
DX_TYPE2	Diagnosis Type #2. Corresponds to Diagnosis #2	72.8	72.1
DX_TYPE3	Diagnosis Type #3. Corresponds to Diagnosis #3	79.7	79.2
DX_TYPE4	Diagnosis Type #4. Corresponds to Diagnosis #4	84.4	84.0
DX_TYPE5	Diagnosis Type #5. Corresponds to Diagnosis #5	87.9	87.6
DX_TYPE6	Diagnosis Type #6. Corresponds to Diagnosis #6	90.5	90.3
DX_TYPE7	Diagnosis Type #7. Corresponds to Diagnosis #7	92.5	92.3
DX_TYPE8	Diagnosis Type #8. Corresponds to Diagnosis #8	94.0	93.9

DX_TYPE9	Diagnosis Type #9. Corresponds to Diagnosis #9	95.2	95.1
DX_TYPE10	Diagnosis Type #10. Corresponds to Diagnosis #10	96.0	95.9
DX_TYPE11	Diagnosis Type #11. Corresponds to Diagnosis #11	96.7	96.6
DX_TYPE12	Diagnosis Type #12. Corresponds to Diagnosis #12	97.2	97.1
DX_TYPE13	Diagnosis Type #13. Corresponds to Diagnosis #13	97.6	97.5
DX_TYPE14	Diagnosis Type #14. Corresponds to Diagnosis #14	97.9	97.8
DX_TYPE15	Diagnosis Type #15. Corresponds to Diagnosis #15	98.1	98.1
DX_TYPE16	Diagnosis Type #16. Corresponds to Diagnosis #16	98.4	98.3
DX_TYPE17	Diagnosis Type #17. Corresponds to Diagnosis #17	98.5	98.5
DX_TYPE18	Diagnosis Type #18. Corresponds to Diagnosis #18	98.7	98.7
DX_TYPE19	Diagnosis Type #19. Corresponds to Diagnosis #19	98.8	98.8
DX_TYPE20	Diagnosis Type #20. Corresponds to Diagnosis #20	99.0	98.9
DX_TYPE21	Diagnosis Type #21. Corresponds to Diagnosis #21	99.1	99.0
DX_TYPE22	Diagnosis Type #22. Corresponds to Diagnosis #22	99.1	99.1
DX_TYPE23	Diagnosis Type #23. Corresponds to Diagnosis #23	99.2	99.2
DX_TYPE24	Diagnosis Type #24. Corresponds to Diagnosis #24	99.3	99.3
DX_TYPE25	Diagnosis Type #25. Corresponds to Diagnosis #25	99.4	99.3
DX_TYPE26	Diagnosis Type #26. Corresponds to Diagnosis #26	99.4	99.4
DX_TYPE27	Diagnosis Type #27. Corresponds to Diagnosis #27	99.5	99.5
DX_TYPE28	Diagnosis Type #28. Corresponds to Diagnosis #28	99.5	99.5
DX_TYPE29	Diagnosis Type #29. Corresponds to Diagnosis #29	99.6	99.5
DX_TYPE30	Diagnosis Type #30. Corresponds to Diagnosis #30	99.6	99.6

¹Denominators exclude health centers with complete missingness for all diagnosis variables (N=91 health centers).

Section 9 Data Comparison

Section 9.1 Public Use Data Files and Restricted Use Data File

Of the 95 participating health centers that were included in the 2023 NAMCS HC Component restricted use data file, all 95 (100%) were selected to create the public use data file sample. The 2023 public use data file contains 450,645 health center visits, for a weighted total of 119,939,235 health center visits (standard error of 16,469,993 health center visits). The 2023 NAMCS HC Component restricted use data file contains unweighted data from the same 95 health centers that submitted 9,012,885 health center visits, for a weighted total of 119,939,021 health center visits (standard error of 16,469,821 health center visits). A comparison of weighted frequencies for health center visits in the public use data file and restricted use data file is presented in Table 9.1.

Table 9.1 Comparison of frequencies for health center visits on the public use data file (weighted n=119,939,235) and restricted use data file (weighted n=119,939,021) for NAMCS HC Component, 2023

Variable	Public Use Data file				Restricted Use Data File			
	Unweighted Count	Weighted			Unweighted Count	Weighted		
		Count	Std. Error	%		Count	Std. Error	%
Age (in years)								
Under 1	9,958	2,647,381	364,118	2.2	204,697	2,711,281	364,773	2.3
1-17 years	75,409	20,070,418	2,651,768	16.7	1,509,802	20,071,364	2,642,325	16.7
18-44 years	160,367	42,503,873	5,659,544	35.4	3,206,451	42,506,801	5,677,970	35.4
45-64 years	133,077	35,422,193	5,050,033	29.5	2,657,732	35,370,313	5,065,097	29.5
65-74 years	47,842	12,928,640	2,243,304	10.8	957,801	12,931,865	2,234,602	10.8
75 years and over	23,986	6,365,345	1,197,455	5.3	476,342	6,346,719	1,195,376	5.3
Missing	6	1,385	630	0.0	60	678	314	0.0
Sex								
Male	164,339	43,702,355	5,898,210	36.4	3,284,427	43,585,683	5,839,816	36.3
Female	282,960	75,013,136	10,546,746	62.5	5,661,069	75,116,382	10,604,967	62.6
Missing	3,346	1,223,744	686,996	1.0	67,389	1,236,956	700,164	1.0
Visit month								
January	38,532	10,433,074	1,453,036	8.7	770,592	10,432,700	1,452,854	8.7
February	38,610	10,307,161	1,411,614	8.6	772,145	10,306,676	1,411,377	8.6
March	41,581	11,368,386	1,580,874	9.5	831,582	11,367,223	1,580,923	9.5
April	35,252	9,673,846	1,371,190	8.1	705,124	9,674,797	1,371,371	8.1
May	40,471	10,954,870	1,574,570	9.1	809,360	10,954,771	1,574,599	9.1
June	39,292	10,560,844	1,521,574	8.8	785,897	10,561,810	1,521,576	8.8
July	34,792	9,060,188	1,291,732	7.6	695,891	9,060,561	1,291,807	7.6
August	40,429	10,606,863	1,534,857	8.8	808,506	10,606,066	1,534,725	8.8
September	35,795	9,352,979	1,392,405	7.8	715,909	9,352,643	1,392,467	7.8
October	37,275	9,673,967	1,466,012	8.1	745,400	9,672,360	1,466,117	8.1
November	38,095	10,016,619	1,447,397	8.4	762,017	10,018,894	1,447,163	8.4
December	30,521	7,930,438	1,210,043	6.6	610,462	7,930,518	1,210,029	6.6
Race								
AIAN	3,527	954,760	212,126	0.8	70,199	947,787	205,902	0.8
Asian	12,381	3,755,839	1,118,976	3.1	248,671	3,756,139	1,109,073	3.1
Black	70,005	16,352,706	4,058,663	13.6	1,395,236	16,290,360	4,044,607	13.6

NHOPI	2,147	776,872	364,698	0.7	42,052	758,871	350,787	0.6
White	210,707	56,685,333	9,274,984	47.3	4,216,418	56,716,104	9,309,259	47.3
Other	25,716	7,247,021	2,042,970	6.0	514,747	7,259,687	2,048,871	6.1
Missing ¹	126,162	34,166,704	9,159,514	28.5	2,525,562	34,210,071	9,167,417	28.5
Ethnicity								
Hispanic or Latino	150,678	43,398,887	9,562,998	36.2	3,009,786	43,343,252	9,568,652	36.1
Not Hispanic or Latino	199,531	49,738,381	7,974,556	41.5	3,991,605	49,748,619	7,977,141	41.5
Missing ¹	100,436	26,801,968	5,667,467	22.4	2,011,494	26,847,150	5,666,902	22.4

¹All health centers, including the health centers with complete missingness in race and ethnicity were included in the comparison of frequencies for race and ethnicity. When presenting analysis, data users should follow the normalization guidance provided in Section 7.2.

Note: All estimates provided in this table do not round to the nearest thousandth for comparison purposes. Data users should round to the nearest thousandth when presenting analyses as indicated in the presentation standards in Section 6.

Section 10 Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA) Checklist for the 2023 NAMCS HC Component

Public Use Data File

Table 10.1 below provides a Preferred Reporting Items for Complex Survey Analysis (PRICSSA) checklist (Seidenberg, Moser, & West, 2023) for users of the 2023 NAMCS HC Component public use data file. This information may be helpful to users when analyzing the 2023 NAMCS HC Component public use data file.

10.1 Preferred Reporting Items for Complex Sample Survey Analysis

Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA)	Description
Name of survey	National Ambulatory Medical Care Survey Health Center Component
Data collection mode	EHR data submission
Target population	Federally qualified health centers (FQHCs) and FQHC look-alikes in the 50 U.S. states and the District of Columbia that used an EHR system
Populations excluded	Health Centers excluded: <ul style="list-style-type: none"> - Indian Health Service Program facilities - Did not provide healthcare services to the general U.S. population, or only provided care to special institutionalized populations such in prisons, nursing homes, homeless shelters, etc. - Only provided dental services - Were located on a military installation or outside of the 50 U.S. states and the District of Columbia
Sample design	Stratified systematic sampling
Variance and standard error estimation	PSU (HCID_S) and Stratum (STRATUM_S) variables were applied and Taylor Series Linearization was used to produce design-adjusted standard errors.
Weighting	VISWT, POPVST
Presentation standards	Proportions or percentages: https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf Rates and counts: https://www.cdc.gov/nchs/data/series/sr_02/sr02_202.pdf
Unweighted total sample size	450,645 visits
Weighted total sample size	119,939,235 visits
Response rate (unweighted)	30.2%
Location of example code	See Section 7

Section 11 Research Data Center

NCHS operates the Research Data Center (RDC) to allow researchers access to restricted-use data. The RDC is responsible for protecting the confidentiality of survey respondents, study subjects, and institutions while providing access to restricted-use data for statistical purposes. The 2023 NAMCS HC Component restricted use data file, which contains unmasked and additional data from all visits at participating health centers (N=9,012,885 visits), can be accessed through the Federal and NCHS RDC. In addition, the 2023 NAMCS HC Component restricted use data file will be linked to other vital and administrative records such as the National Death Index (NDI), U.S. Housing and Urban Development (HUD) administrative data, and others. The linked data will both expand the analytic utility of the NAMCS HC Component data and provide the opportunity to conduct a vast array of studies focused on the associations between a variety of health factors, health care utilization, housing situations, and mortality.

For information on how to access the 2023 NAMCS HC Component restricted use data file through the RDC, please see: <https://www.cdc.gov/rdc/b1datatype/Dt1224a.htm>.

Section 12 References

Health Level Seven International. HL7 CDA® R2 Implementation Guide: National Health Care Surveys (NHCS), R1 STU Release 3.1 - US Realm. Available at: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=385. Accessed November 8, 2024.

Lumley T. “survey: analysis of complex survey samples.” R package version 4.2. 2023. Available at: <https://cran.r-project.org/web/packages/survey/index.html>.

Parker JD, Talih M, Malec DJ, et al. National Center for Health Statistics data presentation standards for proportions. National Center for Health Statistics. Vital Health Stat 2(175). 2017. Available at: https://www.cdc.gov/nchs/data/series/sr_02/sr02_175.pdf

Parker JD, Talih M, Irimata KE, Zhang G, Branum AM, Davis D et al. National Center for Health Statistics data presentation standards for rates and counts. National Center for Health Statistics. Vital Health Stat 2(200). 2023. DOI: <https://dx.doi.org/10.15620/cdc:124368>.

Seidenberg AB, Moser RP, West BT. Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA). J Surv Stat Methodol 11(4). 2023. <https://doi.org/10.1093/jssam/smac040>.

StataCorp. Stata 18 Survey Data Reference Manual. College Station, TX: Stata Press. 2023. Available at: <https://www.stata.com/manuals/svsvyset.pdf>.

Williams SN, Ukaigwe J, Ward BW, Okeyode T, Shimizu IM. Sampling procedures for the collection of electronic health record data from federally qualified health centers, 2021-2022 National Ambulatory Medical Care Survey. National Center for Health Statistics. Vital Health Stat Series 2(203). 2023. DOI: <https://dx.doi.org/10.15620/cdc:127730>.

Appendix A Unweighted frequencies for health center visits

Appendix Table A.1. Unweighted frequencies for health center visits on the public use data file, National Ambulatory Medical Care Survey Health Center Component, 2023 (n=450,645)

Variable	Description	Count	%
YEAR	Survey year		
2023		450,645	100.0
DAY	Day of the week		
1	Sunday	3,155	0.7
2	Monday	86,732	19.3
3	Tuesday	94,947	21.1
4	Wednesday	92,497	20.5
5	Thursday	89,845	19.9
6	Friday	75,376	16.7
7	Saturday	8,093	1.8
MONTH	Month of visit		
1	January	38,532	8.6
2	February	38,610	8.6
3	March	41,581	9.2
4	April	35,252	7.8
5	May	40,471	9.0
6	June	39,292	8.7
7	July	34,792	7.7
8	August	40,429	9.0
9	September	35,795	7.9
10	October	37,275	8.3
11	November	38,095	8.5
12	December	30,521	6.8
MARITAL	Marital status		
-9	Missing	117,868	26.2
D	Divorced	21,920	4.9
L	Legally Separated	5,562	1.2
M	Married	96,007	21.3
O	Other	42	0.0
S	Single	80,176	17.8
T	Domestic Partner	5,555	1.2
U	Unmarried	111,133	24.7
W	Widowed	12,382	2.8
AGE_GROUP	Patient age group		
-9	Missing	6	0.0
1	Less than 18 years	85,367	18.9
2	18-44 years	160,367	35.6
3	45-64 years	133,077	29.5
4	65 years or more	71,828	15.9
ETHNICITY	Patient Hispanic ethnicity		
-9	Missing	100,436	22.3
1	Hispanic or Latino	150,678	33.4

2	Not Hispanic or Latino	199,531	44.3
RACE	Patient race		
-9	Missing	126,162	28.0
1	AIAN	3,527	0.8
2	Asian	12,381	2.8
3	Black	70,005	15.5
4	NHOPI	2,147	0.5
5	White	210,707	46.8
6	Other	25,716	5.7
RACERETH	Patient race and Hispanic ethnicity		
-9	Missing	86,455	19.2
1	White	124,207	27.6
2	Black	64,193	14.2
3	Hispanic	150,678	33.4
4	Other	25,112	5.6
SEX	Patient sex		
-9	Missing	3,346	0.74
1	Male	164,339	36.5
2	Female	282,960	62.8

Appendix B Health Center (HC) Facility Interview

Form Approved
OMB No. 0920-0234
Exp. date 11/30/2025

Notice – CDC estimates the average public reporting burden for this collection of information as 45 minutes per response, including the time for reviewing instructions, searching existing data/information sources, gathering and maintaining the data/information needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to CDC/ATSDR Information Collection Review Office, 1600 Clifton Road, MS D-74, Atlanta, GA 30333; ATTN: PRA (0920-0234).

Assurance of confidentiality – We take your privacy very seriously. All information that relates to or describes identifiable characteristics of individuals, a practice, or an establishment will be used only for statistical purposes. NCHS staff, contractors, and agents will not disclose or release responses in identifiable form without the consent of the individual or establishment in accordance with section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)) and the Confidential Information Protection and Statistical Efficiency Act of 2018 (CIPSEA Pub. L. No. 115-435, 132 Stat. 5529 § 302). In accordance with CIPSEA, every NCHS employee, contractor, and agent has taken an oath and is subject to a jail term of up to five years, a fine of up to \$250,000, or both if he or she willfully discloses ANY identifiable information about you. In addition to the above cited laws, NCHS complies with the Federal Cybersecurity Enhancement Act of 2015 (6 U.S.C. §§ 151 and 151 note) which protects Federal information systems from cybersecurity risks by screening their networks.

Name: _____
Title: _____
Contact number: _____
Extension: _____
Call notes: _____

Initial Confirmation and Telephone Screen Call

1. Will you please tell me if the following information is correct?

Health center name: _____
Authorized Official Name and Salutation: (Mr./Ms./Miss/Mrs./Dr.) _____
Authorized Official Title: _____
Email Address: _____
Address: _____
City/State/ZIP code: _____
Telephone number: _____ Extension: _____

2. Which of the following best describes your center?

- ☐ Federally Qualified Health Center (330 grantee) → *CONTINUE WITH Q3*
- ☐ Federally Qualified Health Center Look-Alike → *SKIP TO Q5*
- ☐ Urban Indian (437) Health Center → *CONCLUDE INTERVIEW*
- ☐ Other (Please Specify) → *SKIP TO Q4*

3. Can your center also be classified as a: *Note: Select all that apply.*

- ☐ Migrant Health Center (MHC)
- ☐ Health Care for the Homeless (HCH)
- ☐ Public Housing Primary Care (PHPC) Grant Program

- ☐ None of the above
- ☐ Blank → *SKIP TO Q5*

4. Other – please Specify: _____

- ☐ Blank → *CONCLUDE INTERVIEW*

5. Are you the official who can agree to participate in NAMCS on behalf of the [insert health center name]?

- ☐ Yes → *SKIP TO Q12*
- ☐ No → *CONTINUE WITH Q6*

6. Will you please identify an official who can agree to participate in NAMCS on behalf of the [insert health center name]?

Salutation/Name: (Mr./Ms./Miss/Mrs./Dr.)

Title:

Telephone number: _____ Extension: _____

Email Address: _____

Address: _____

City/State/ZIP Code: _____

7. Can you please confirm if [insert title from Q1 or Q6 or Q11] received an information packet and invitation to participate in NAMCS?

- ☐ Yes
- ☐ No

8. Can you please transfer me to [insert name from Q1 or Q6 or Q11]?

- ☐ Yes → *SKIP TO Q10*
- ☐ No → *CONTINUE WITH Q9*

9. When would be a good time to call back and speak with the [insert title from Q1 or Q6 or Q11]?

_____/_____/_____
Day / Month / Year

Time: ____:____ A.M. ____ P.M. ____ Time Zone

Call Note: _____

- ☐ Blank

10. As the [insert title from Q1 or Q6 or Q11], are you *authorized* to agree to participate on behalf of [insert health center name]?

- ☐ Yes → *SKIP TO Q12*
- ☐ No → *CONTINUE WITH Q11*

11. Who is the best person who can *authorize* participation in the survey?

Name: (Mr./Ms./Miss/Mrs./Dr.)

Job title:

Telephone Number: _____ Extension: _____

E-mail: _____

Address: _____

City/State/ZIP code: *GO BACK TO Q7*

Interview with Health Center Official

12. Did you receive the NAMCS information packet?

- ☐ Yes, no need to send it again → *SKIP TO Q17*
☐ Yes, but can you please send it again → *CONTINUE WITH Q13*
☐ No → *CONTINUE WITH Q13*

13. Would you prefer to receive it via email or in the mail?

- ☐ Email → *CONTINUE WITH Q14 TO CAPTURE AND SEND EMAIL*
☐ Mail → *SKIP TO Q15 TO CONFIRM MAILING ADDRESS TO BE USED TO SEND A NEW LETTER*
☐ Blank → *CONTINUE WITH Q14*

14. What email address would you like us to send it to? _____ *SKIP TO Q16*

- ☐ Blank → *CONTINUE WITH Q15*

15. Could you please confirm the following contact information?

Name: (Mr./Ms./Miss/Mrs./Dr.) _____

Health Center name: _____

Address: _____

City, State and ZIP code: _____

E-mail: _____

- ☐ Valid
☐ Invalid

16. When would be a good time to call back?

_____/_____/_____
 Day / Month / Year

Time: _____ : _____ A.M. _____ P.M. _____ Time Zone

Call Note: _____

- ☐ Blank

17. Do you have any questions about the information you received or concerns about what we have discussed so far?

- ☐ Yes → *CONTINUE WITH Q18*
☐ No → *SKIP TO Q19*
☐ Blank → *SKIP TO Q19*

18. Record major topics below. Use materials to try to address each one.

- ☐ Blank _____

19. Can we count on your health center's participation in NAMCS?

- ☐ Yes → *CONCLUDE INTERVIEW*
☐ Need more time to decide → *CONTINUE WITH Q20*
☐ No, health center official declines to participate. → *SKIP TO Q21*

20. When would be a good time to call back?

_____/_____/_____
 Day / Month / Year

Time: _____:_____ A.M. _____ P.M. _____ Time Zone

Call Note: _____

☐ Blank

CONCLUDE INTERVIEW. DURING CALL BACK, GO BACK TO Q19

21. Please tell me why your health center does not want to participate.

RECORD RESPONSE TO BE CODED LATER: _____

☐ Blank

22. DO NOT READ THESE RESPONSES OUT LOUD; Instead; check the option that best captures the official's reason for refusal.

☐ Confidentiality concerns → *CONCLUDE INTERVIEW*

☐ The health center's financial situation does not permit it to dedicate time to this effort → *CONCLUDE INTERVIEW*

☐ The health center has too many other priorities at this time → *CONCLUDE INTERVIEW*

☐ Limited staffing resources → *CONCLUDE INTERVIEW*

☐ Other → *CONTINUE WITH Q23*

☐ Blank → *CONCLUDE INTERVIEW*

23. Other – please specify: _____

☐ Blank

Health Center Primary Contact Interview

24. Is this health center a subsidiary of a larger company or network?

☐ Yes → *CONTINUE WITH Q25*

☐ No → *SKIP TO Q26*

☐ Don't know → *SKIP TO Q26*

25. What is the name of the larger company or network?

☐ Blank

26. Are other health centers covered under your state license?

☐ Yes → *CONTINUE WITH Q27*

☐ No → *SKIP TO Q28*

☐ Don't know → *SKIP TO Q28*

☐ Blank → *SKIP TO Q28*

27. What are the name(s) of the health center(s)?

☐ Blank

28. When this health center reports data to the governing bodies is the information solely for this health center or are other health centers included in the data transmission?

☐ Solely for this health center → *SKIP TO Q30*

☐ Combined with one or more other health centers → *CONTINUE WITH Q29*

☐ Blank → *SKIP TO Q30*

29. What are the name(s) of the other health centers?

Note: Health center care delivery sites under the sampled health center with the same name or a different name are not considered as "other health centers."

☐ Blank

Part 2. General Questions

30. Was this health center open for the full calendar year (FILL PREVIOUS CALENDAR YEAR)?

☐ Yes → *SKIP TO Q32*

☐ No → *CONTINUE WITH Q31*

☐ Never open in (FILL PREVIOUS CALENDAR YEAR) → *SKIP TO Q32*

31. Please provide the dates the health center was closed in (FILL PREVIOUS CALENDAR YEAR): _____

PERIOD 1: ☐ Exact ☐ Estimate

PERIOD 2: ☐ Exact ☐ Estimate

PERIOD 3: ☐ Exact ☐ Estimate

32. Do you anticipate any significant changes in your visit volume in (FILL CURRENT CALENDAR YEAR)?

☐ Yes → *CONTINUE WITH Q33*

☐ No → *SKIP TO Q34*

33. Please explain: _____

☐ Blank

34. During its last normal year, approximately how many office visit encounters did this health center have?

☐ Enter number of visits: _____

35. Approximately how many office visit encounters do you estimate this health center will have in (FILL CURRENT CALENDAR YEAR)?

☐ Enter estimated visits: _____

36. Please provide the actual counts or your best estimates for the total number of health center visits during calendar year (FILL PREVIOUS CALENDAR YEAR) for each quarter if possible, and for the year overall.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4	Annual
All visits made to health center:	<input type="checkbox"/> Blank	<input type="checkbox"/> Blank	<input type="checkbox"/> Blank	<input type="checkbox"/> Blank	

Electronic Health Records (EHR)

37. Are you able to electronically output patient level data from your electronic health record (EHR) system?

- ☐ Yes
☐ No
☐ Don't know

38. What is your health center's EHR vendor, product and version number?

EHR Vendor: _____

EHR Product: _____

EHR Version Number: _____

39. Does your health center already have an established Health Information Service Provider (HISP) that can be used to transfer secure Health Insurance Portability and Accountability Act of 1996 (HIPAA) compliant data?

- ☐ Yes
☐ No
☐ Don't know

40. Do you need assistance setting up your EHR system to ensure that it is compatible with the Health Level 7 (HL7) Clinical Document Architecture (CDA®) R2 Implementation Guide (IG): National Health Care Surveys (NHCS) Release 1, Draft Standard for Trial Use (DSTU) 1.2 – U.S. Realm (NHCS IG 1.2)?

- ☐ Yes
☐ No
☐ Don't know

41. Will the data you provide include electronic health records from your health center only?

- ☐ Yes → *SKIP TO Q45*
☐ No → *CONTINUE WITH Q42*
☐ Don't know → *SKIP TO Q45*

42. Is it possible to identify the records from your health center separate from the other health centers that report with you?

- ☐ Yes
☐ No
☐ Don't know

Data Transfer

43. What are the name(s) of the other health centers included?

44. How can we make that distinction? _____

45. Will you please identify the primary IT/data contact, the individual responsible for transmitting your health center's data and what is their contact information?

Name: (Mr./Ms./Miss/Mrs./Dr.)

Job title:

Telephone Number: Extension:

E-mail:

COVID-19 Information**46. Does your health center offer COVID-19 vaccinations?**☐ Yes → *CONTINUE WITH Q47*☐ No → *SKIP TO Q49*☐ Blank → *SKIP TO Q49***47. Which vaccine does your health center offer? SELECT ALL THAT APPLY.**☐ Moderna → *SKIP TO Q49*☐ Johnson & Johnson/Janssen → *SKIP TO Q49*☐ Pfizer → *SKIP TO Q49*☐ Other → *CONTINUE WITH Q48*☐ Don't know → *SKIP TO Q49*☐ Blank → *SKIP TO Q49***48. Other – please specify: _____****Electronic Health Records (EHR) and Telemedicine**

49. Does your health center use an EHR to...?	Yes	No	Don't Know	Not Applicable	Blank
Record social determinants of health (e.g., employment, education, race/ethnicity, language and literacy skills)?					
Record behavioral determinants of health (e.g., tobacco use, physical activity, alcohol use, drug use, diet)?					
Order prescriptions?					
Send prescriptions electronically to the pharmacy?					

50. At your health center, what type(s) of telemedicine do you use for patient visits?*Note: Check all that apply.*☐ Videoconference software with audio (e.g., Zoom, WebEx, FaceTime) → *SKIP TO Q52*☐ Audio without video conference software → *SKIP TO Q52*☐ Telemedicine platform NOT integrated with EHR (e.g., Doxy.me) → *SKIP TO Q52*☐ Telemedicine platform integrated with EHR (e.g., update clinical documentation during telemedicine visit) → *SKIP TO Q52*☐ Other tool(s) → *CONTINUE WITH Q51*☐ The health center does not use telemedicine for patient visits → *SKIP TO Q53*☐ Blank → *SKIP TO Q52*

51. Other tool(s) – please specify: _____ → *CONTINUE*
WITH Q52

52. At your health center, in a typical week, how many of your visits use telemedicine?

- ☐ None
- ☐ Some
- ☐ Most
- ☐ All
- ☐ Blank

53. At your health center, what, if any, issues affect your use of telemedicine?

Note: Select all that apply.

- ☐ Limited Internet access and/or speed issues
- ☐ Telemedicine platform not easy to use
- ☐ Telemedicine isn't appropriate for my specialty/type of patients
- ☐ Limitations in patients' access to technology (e.g., smartphone, computer, tablet, Internet)
- ☐ Patients' difficulty using technology/telemedicine platform
- ☐ Improved reimbursement and relaxation of rules related to use of telemedicine visits
- ☐ Blank

Payment Information

54. Please provide the following information to indicate to whom the \$10,000 set-up fee checks should be mailed to.

Payee (Health Center Name):

Payee Point of Contact Name (Mr./Ms./Miss/Mrs./Dr.):

Attn:

Job Title:

Address:

City/State/ZIP Code:

Telephone Number:

Extension:

E-mail: