



National Health &
Aging Trends Study

how daily life changes as we age

nSOC



National Study of Caregiving

NHATS Restricted Data Vetting Handbook

NHATS Restricted Data Committee

Table of Contents

SUMMARY	3
I. RESPONSIBILITY TO PROTECT INFORMATION	4
II. GENERAL GUIDELINES FOR PREPARING A VETTING REQUEST FOR AN EXPORT	4
I. THE MINIMUM CELL COUNT RULE	4
II. SELF-VETTING EXPORT REQUESTS	5
III. LIMITS ON FILES ALLOWED OUT OF THE ENCLAVE	6
IV. MORE SPECIFIC VETTING RULES FOR TYPES OF EXPORT REQUESTS	6
I. DESCRIPTIVE TABLES	6
II. MODEL BASED ESTIMATES	7
III. SURVIVAL ANALYSES	7
V. GEOGRAPHIC ANALYSES	8
I. GEOGRAPHIC ANALYSES WITH NHATS (MICDA ENCLAVE)	8
II. GEOGRAPHIC ANALYSES WITH NHATS-CMS LINKED FILES (LINKAGE ENCLAVE)	8
VI. PROVIDER AND PLAN ANALYSES WITH NHATS-CMS LINKED FILES (LINKAGE ENCLAVE)	8
VII. GRAPHICS	8
VIII. OTHER CONSIDERATIONS	9
I. MISSING VALUES AND IMPUTATION	9
II. THE USE OF WEIGHTS	9
III. RESIDUAL DISCLOSURE RISK	9
IX. NHATS GUIDELINES FOR THE RELEASE OF CODE	10
X. PROCESS WHEN AN EXPORT DOES NOT MEET RELEASE CRITERIA	10
XI. LIMITS ON FILES REQUESTED FOR IMPORT	11
XII. PROCEDURE FOR REQUESTING AN IMPORT TO OR EXPORT FROM THE ENCLAVE	11
II. MICDA ENCLAVE IMPORT REQUESTS	11
I. MICDA ENCLAVE EXPORT REQUESTS	11
IV. LINKAGE ENCLAVE IMPORT REQUESTS	12
III. LINKAGE ENCLAVE EXPORT REQUESTS	12
APPENDIX A: VETTING REQUEST FORM (NOT PUBLICLY AVAILABLE)	14
APPENDIX B: LIST OF COMMON ISSUES IN EXPORT REQUESTS THAT VIOLATE VETTING RULES	15
APPENDIX C: REFERENCES	17

Summary¹

This guide is designed to outline considerations for protecting confidential data contained in the National Health and Aging Trends Study (NHATS) and the National Study of Caregivers (NSOC), hereafter referred to collectively as NHATS, as well as their linkages to existing datasets from the Centers for Medicare & Medicaid Services (CMS).

Access to NHATS restricted files (including Genetics files) is limited to users who meet the requirements of NHATS and the Michigan Center on the Demography of Aging (MiCDA) enclave. These data are accessed through the MiCDA enclave managed by the University of Michigan.

Access to NHATS-CMS Standard Linked files (hereafter Standard files) and NHATS-CMS Provider Linked files (hereafter Provider files) are limited to users who meet the requirements of NHATS and the National Institute on Aging (NIA). The linked data are accessed through the NIA Data LINKAGE Program (LINKAGE) enclave.

Before any output can be released from the MiCDA or LINKAGE enclaves, it must be reviewed and approved for release by an NHATS staff member who has the proper training and responsibility for reviewing and releasing output. The primary goal of this disclosure avoidance review (hereafter “vetting”) process is to protect the privacy of survey respondents and the confidentiality of the data that they provide. To minimize the risks of disclosure (or identification) of respondents or the areas in which they live, as well as providers that serve them, NHATS places limits on the files that are allowed both in and out of the enclaves.

Researchers and NHATS staff both play a critical role in protecting the confidentiality of NHATS data. As stewards of the data, NHATS staff retain final authority over what is allowed to be brought in or taken out of the enclave.

This document is updated regularly by the NHATS Restricted Data Committee. Researchers are encouraged to ensure that they have the most current version by visiting:
<https://nhatsannex.org/nhats-nsoc-restricted-data>

¹ These guidelines were informed by Lauger et al. (2014), Hawes et al. (2025), Statistics Canada Data Access Division Vetting Committee (2022), and the U.S. Census Bureau Disclosure Review Board (2023).

I. Responsibility to Protect Information

Researchers and NHATS staff work together to (1) ensure that all researchers are aware of the vetting rules for the datasets (and also of any changes to the vetting rules that may impact their project); (2) review and monitor output to ensure that the vetting rules are satisfied; and (3) ensure that files requested for import into the enclave satisfy vetting rules.

In the event that the vetting rules are not satisfied, NHATS staff aim to work with researchers to find solutions whenever possible.

The **primary responsibility of researchers** is to ensure that the applicable vetting rules are adhered to so NHATS participants cannot be identified based on the output released from the enclave. Often this can be achieved by following a few simple rules (see Appendices A & B). It is the researcher's responsibility to provide clear documentation showing that all rules have been followed and to self-vet their output before requesting review from NHATS staff. The documentation should provide support files that detail: relevant statistics or tabulations, variable and sample definitions, and, if applicable, information on how prior vetting requests relate to the current request. Some general questions for researchers to consider when preparing outputs for vetting include:

- Is the output consistent with what was written in the approved research plan?
- Can the output be understood by someone who is not on the research team? Is all output clearly labelled?
- Has the output been checked thoroughly for errors?
- Is **all** the requested output **really needed** outside of the enclave?
- Has the primary investigator or research team agreed that the output is ready to be submitted for vetting?
- Are changes in sample composition anticipated? Is this the final version of the output? Can the descriptive statistics be released at **the end** of the project?

Once output is released from the enclave, it is considered public information. All researchers that have access to restricted data sign a data use agreement (DUA) with NHATS and Johns Hopkins University that requires them to uphold data confidentiality. The disclosure avoidance methods used by NHATS are specific to NHATS and may not mirror those of other datasets or institutions. The disclosure minimization rules are designed to protect participants and ensure that the greater public can trust that any information gathered by researchers is not disclosed without their permission. The rules are also designed so that they may be applied consistently by different staff across different projects.

II. General Guidelines for Preparing a Vetting Request for an Export

To facilitate self-vetting and vetting by the NHATS team, researchers requesting information to be exported from the enclave must provide clearly labeled, understandable output in tabular form, with clearly labeled supporting information.

I. The Minimum Cell Count Rule

Each statistic, coefficient, number, etc. that is requested to be released has an underlying sample of original person observations in the NHATS datasets. That underlying sample must pass all relevant disclosure rules, as documented in supporting statistics. The primary disclosure rule is that individual

observations may never be identified from the results; to achieve this, **all underlying samples for results that are exported must contain 11 or more observations²**.

Researchers must provide supporting statistics for all created samples that meet this minimum cell count before any information is released from the enclave. Typically, these supporting (or disclosure) statistics are tabulations of the sample sizes of all created samples. Such statistics include sample sizes for all variables and cell counts for categorical/binary variables. These supporting statistics should:

- show that the requested output passes vetting rules for all requested output, **for each sample requested**
- be unweighted (even if statistics requested to be removed are weighted)
- document implicit samples, which are samples of categories or missing observations that can be calculated through arithmetic

Additional supporting output may be required, depending on the request (see III. Limits on Files Allowed Out of the Enclave and Appendix B).

II. Self-Vetting Export Requests

Self-vetting, or the process of researchers examining their own results for compliance with NHATS disclosure minimization rules, is a crucial part of the vetting process. At a minimum, results that are approved for export from the enclave must make the sample and type of analysis clear. Once researchers have answered the general questions for consideration (listed in I. Responsibility to Protect Information), questions that researchers should ask before submitting results for export are:

- Is the name of the dataset(s) that is/are the basis of the results clearly identified?
- Are the sample sizes for each table or analysis noted clearly?
- Does each table or analysis have a clear description of the sample or sub-sample and the population or sub-population for each regression or table?
- Are there clear labels for all variables or a glossary of variables included in the export request?
- Is the dependent variable clearly identified?
- Are categorical variables clearly identified? Is there supporting documentation noting the cell sizes for each categorical variable? If the analysis has a categorical dependent variable, is there supporting documentation noting the cell sizes for each categorical covariate by the dependent variable?
- Are the results weighted and if so, is this noted on each table or analysis?

Other general rules and guidelines that researchers should note are:

- Only analyses at the individual level are allowed for release. Other samples, such as geographic units (e.g., tracts, counties, states, HRRs), facilities (hospitals, nursing homes, assisted living), or providers are not allowed to be the unit of analysis.
- Program files (e.g., code) must not contain notes about sample sizes.
- No hidden information is allowed in files requested for export from the enclave.
- Empty cells that have been suppressed due to small sample sizes are not allowed; researchers must combine rows or columns so that the small cell is not evident.

² This is a disclosure minimization rule also mandated by CMS for any projects that use restricted CMS data.

As a general rule, researchers should not be reporting in publications or presentations or otherwise sharing that a cell size for a particular group is less than 11. Researchers will be asked to collapse cells if a cell size is less than 11.

III. Limits on Files Allowed Out of the Enclave

Information may **only** be removed from the enclave after being reviewed by NHATS staff. **Transcription, screenshots, screensharing, and other forms of information transfer that circumvents the disclosure review process are not allowed.**

Prior to removal, all files will be reviewed for disclosure risk by NHATS (see below for a list of general rules). **Each request to export files out of the enclave must be accompanied by a vetting request form (Appendix A).**

Files **allowed** out of the enclave, pending review:

- Well-labelled tables and graphs that meet the vetting requirements described below. Files will be prioritized and typically reviewed within 5 business days, if possible.
- Statistical code (e.g., .do or .sas files) without notes about sample sizes.
- Log/output files (e.g., .txt or .scml) are allowed to be removed from the enclave in **exceptional circumstances** only. Strong justification is needed for such requests. Such files will take lower priority and may take longer than average to review.

Files **not allowed** out of the enclave:

- Microdata files
- Geographic visualizations (maps)
- Graphs and other visualizations that show individual observations (e.g., sequence analysis, scatterplots, boxplots with outliers)

IV. More Specific Vetting Rules for Types of Export Requests

Details on more specific disclosure minimization rules for the most common types of analysis are below. This is not an exhaustive list of rules related to analyses or information requested for export.

I. Descriptive Tables

Descriptive statistics are among the most sensitive results NHATS allows for export from the enclave. **NHATS requests that researchers limit their tabular output to the minimum necessary.** Below are our general rules that apply to descriptive tables requested for export:

- All unweighted table cell counts must be provided in tables requested for export **or** accompany weighted tables as supporting statistics
- All unweighted table cell counts must be at least 11, including counts that can be obtained through arithmetic
- All item missingness should be included in the table; missing observations should be combined with another cell (imputed or recoded) if $n < 11$
- All categories of a variable must be included or collapsed
 - NHATS does not allow cell suppression (i.e., dropping a category because it is too small)

- Redaction of a single cell, column or row is not acceptable (the omitted group is known to be small by omission)
- Text may not report which specific categories/values have been imputed or combined because of small sample size
- Minima and maxima (e.g., top and bottom of ranges for variables) must have at least 11 cases with top and bottom values if shown
- Unweighted table cell counts should sum to total analytic sample size
- For means of binary variables, numerators and denominators must be shown
- For means of binary variables, numerators must be at least 11 (and sample n minus numerator must be at least 11)
- Specific geographic areas below Census division (e.g., Middle Atlantic, Pacific, etc.) may not be reported
- Output may not list or print cases. Microdata is not releasable.
- Categorical contextual data must have accompanied supporting documentation showing at least 3 states or 11 units for lower levels of geography per category

II. Model Based Estimates

The underlying samples for model based estimates that are requested for export must contain 11 or more observations.

- All covariates must have an unweighted minimum cell size of 11 cases; these cell counts must be submitted as supporting documentation with every model export
- In a categorical dependent variable model, a crosstab of the predictor variable by the outcome is required as supporting documentation and must contain cells of 11 or more
- Minima and maxima (e.g., top and bottom of ranges for variables) must have at least 11 cases if shown
- Predicted values of a variable must have cell sizes of at least 11
- Multilevel models with individuals nested in places may not include coefficients of individuals or places
- Categorical contextual data must have accompanied supporting documentation showing at least 3 states or 11 units for lower levels of geography per category

III. Survival Analyses

Any actuarial survival function, such as a life table or a Kaplan-Meier analysis, can be released when: the number of events per interval or cycle and the number of units remaining are both either zero or at least 11. Censoring counts due to nonresponse or lost contact do not need to meet any minimum thresholds and are not problematic for confidentiality vetting. When creating intervals, the intervals do not need to be of equal interval lengths, and intervals can be collapsed to generate a releasable life table. For any survival analysis function-building method (life table, product-moment, or related method like Breslow, etc.), the coefficients (betas) or the parameters are releasable if the model is a releasable model – i.e., not equivalent to an unreleaseable table.

Any life table which does not group events in intervals or survivors at the end of the observation period to respect the cell size minimum is not releasable.

Any survival or hazard curve which has distinct steps, inflections, gaps, dots or marks at each event and/or censoring is not releasable unless it is based on a releasable table.

V. Geographic Analyses

I. Geographic Analyses with NHATS (MiCDA enclave)

Tabulations, visual representations, or coefficients based on geographic areas below the Census Division may not be removed from the enclave.

Categorical contextual variables (e.g., county expenditures in categories) have additional limitations. Researchers must make sure there are at least 3 states and 11 units for lower levels of geography (in this example, counties) in each category. Each export request requires supporting tabulations of categorical contextual variables at the geographic unit (e.g. a tabulation of Census Tracts in each category of the variable).

Models may not identify individual states or other geographic entities (e.g., county, zip code, etc.). For example, fixed effect coefficients cannot be reported at the state level. Implicit state-level comparisons (e.g., where one state is left out of a model to check for robustness) are not allowed for export.

II. Geographic Analyses with NHATS-CMS Linked Files (LINKAGE enclave)

Geographic analyses with linked NHATS-CMS files are generally not approved as they raised disclosure risk significantly.

Information on Hospital Referral Region (HRR), Census division, and dichotomous metro/non-metro area is available in the LINKAGE enclave for use with **Standard** files. No contextual data (e.g., area deprivation index) may be linked at this time to HRR.

NHATS does not currently approve projects to use Hospital Referral Region (HRR) data or State information embedded in provider variables with **Provider** files. At this time, only Census division and dichotomous metro/non-metro area may be used in conjunction with the Provider Linked files.

Geographic information such as state that may be embedded in the provider ID may not be used as a control variable or otherwise identified in any results.

VI. Provider and Plan Analyses with NHATS-CMS Linked Files (LINKAGE enclave)

Categorical provider-level or plan-level (contextual) variables (e.g., characteristics of providers such as profit status or size in categories; characteristics of plans such as coverage or benefit limits) have additional limitations.

Provider/plan contextual data must be de-identified before being brought into the enclave.

Researchers must make sure there are at least 11 providers/plans in each category. This rule may require additional tabulations of approved contextual data.

Tabulations and/or visual representations or coefficients based on provider or plan may not be removed from the enclave.

VII. Graphics

Graphical output may be released in .gph, .tiff, .png, .svg, and .pdf formats. Graphs shall not contain data or be reversible into data. Other rules regarding charts and plots are:

- Graphical depictions of individual observations are not allowed for release
- Scatterplots are not allowed
- Histograms require supporting documentation that each bar has at least 11 cases
- Line charts must meet frequency requirements at each point
- Categorical contextual data must have accompanied supporting documentation showing at least 3 states or 11 units for lower levels of geography per category
- Outliers cannot be shown in a box and whisker plot/boxplot
- Maps created with NHATS data in the enclave are never allowed outside of the enclave

VIII. Other Considerations

I. Missing Values and Imputation

Almost all databases or administrative records have missing values, and NHATS is no exception. NHATS generally has very low missing data and researchers interested in imputation should familiarize themselves with the NHATS interview guide to avoid imputing valid skips.

Cell counts of missing values that can be calculated through arithmetic that do not meet the minimum cell count rule will not be allowed for export. For descriptive statistics that have missing values cell counts under 11 (but not zero), researchers should collapse rows or columns, revise the table to reflect the final analytic sample, or impute missing cases.

For analyses using imputed data, sample-size totals from the original (i.e., un-imputed) and final imputed dataset may be released as long as each table follows minimum cell count rules. The number of units imputed for individual variables should not be released. Researchers can summarize the range of imputed cases across variables (e.g., 1%-3%), but may not identify if a specific variable had imputed cases under the cell count limit (i.e., < 11).

II. The Use of Weights

NHATS comes with a set of base weights, tracker weights, and analytic weights that are used to produce population-level estimates. Weights can play a role in protecting data confidentiality, but weighting is not required for the release of the estimates derived from NHATS restricted data. Vetting focuses on unweighted estimates. Having weighted counts that meet minimum cell requirements is not a relevant consideration.

III. Residual Disclosure Risk

Descriptive statistics (e.g., counts, means, percentages) pose the greatest risk of identification of NHATS respondents. Minor adjustments to the sample or to variables can create situations where a small number of individuals change categories and as a result become identifiable through a combination of descriptive statistics (this is called residual disclosure). **NHATS encourages researchers to only export descriptive statistics that are the final versions.**

Prior releases will be checked whenever possible for residual disclosure risk. If NHATS staff finds differences between the current export request and prior released versions that identify cells under 11, additional procedures apply.

IX. NHATS Guidelines for the Release of Code³

Most code created by researchers may be released after being vetted by NHATS staff; however, code for public release should not explicitly contain or implicitly reference any sensitive or potentially sensitive information. Code that is requested for release outside of the secure data enclave should read as if it were written **outside** of the enclave. Below are some guidelines that must be followed before code is released.

- Neither the code itself nor the comments may reveal any information about the underlying data, regardless of whether the information revealed falls under usual definitions of PII. All such information must be redacted or go through the normal output review process. If a special case such as an outlier requires special treatment, researchers should take extra care that the part of the code referring to the special case does not reveal anything about a small number of records.
- Code or programs that explicitly include or implicitly describe counts, sample sizes, estimates, etc. will be fully vetted for disclosure risks similar to any other results export request. In such cases, supplementary tables showing minimum cell sizes will be required.
- Code cannot reveal the geographic/contextual information contained in the data.
- Code comments can be problematic if they contain references to the sample, e.g., “Only a few observations were dropped here”, or “No values in sample A changed in this code block” or “4 cases dropped through linking.” These comments will not be allowed outside the enclave and prior to code export requests, researchers are encouraged to removed comments such as these from files.
- Code cannot include identifier values (e.g., SPID = 12345 or ID = 54321).
- For projects in the LINKAGE enclave, we follow additional rules set by NIA, including: “Exportation of code is limited to 10 times per year or during the closeout of the study - all code exports must exclude file paths and PII/PHI before submission.”⁴

X. Process When an Export Does Not Meet Release Criteria

NHATS Repository staff, in consultation with the NHATS PIs, have the final decision over whether a given set of results may be exported. All export requests must document that all relevant disclosure requirements are met. This includes clear supporting documentation, clearly labeled output, as well as providing all required disclosure statistics.

If NHATS staff find that output does not meet the criteria for release, the researcher will be asked to either: 1) collapse/combine cells or 2) reconsider the output by using alternative ways to convey information, such as summarization tables. NHATS staff will communicate directly with the research team member over email regarding exports that do not meet release criteria; whenever possible NHATS staff will CC the team’s primary investigator for transparency.

³ Borrowed substantially from U.S. Census (2023).

⁴ See: <https://www.nia.nih.gov/research/dbsr/nia-data-linkage-program-linkage/disclosure-review-and-file-transfer-processes#file-export-request-process>

XI. Limits on Files Requested for Import

Documents and files requested for import to the enclave will be vetted by NHATS staff. The following files are allowed, pending review:

1. Non-identifiable datasets, aggregated at a geographic or other categorical level, listed on the requesting research team's NHATS restricted data application File Request Form and described in their research plan; and
2. Statistical code (e.g., .do or .sas files) and other supporting documents. All import requests must include a description of the file and justification of need.

XII. Procedure for Requesting an Import to or Export from the Enclave

II. MiCDA Enclave Import Requests

- Place the file(s) for review and a text file describing the file in a folder labeled "Import-MM-DD-YYYY" in your SFTP space.
- Submit a request to nhats-restricted-data@umich.edu.

NHATS Staff will review your list of approved files to make sure the file is listed and then vet the import file(s) for content. Once approved, the requested files will be moved to the requester's folder. Once the process is complete, the researcher will receive notification via email.

If the file (excluding statistical code) is not listed in your research plan and external file request form, you will be asked to update your list of approved files in a new external file request form. If the file does not directly correspond to your research plan, you will be asked to provide a revised research plan, which will have to be approved before the file can be imported.

I. MiCDA Enclave Export Requests

Once self-vetting is complete:

- Place the vetting request form, the file(s) for review, along with supporting documentation in a folder labeled "Export-MM-DD-YYY" in your personal storage space (U: drive).
- Submit a request to nhats-restricted-data@umich.edu.
- A reviewer will vet the files to be exported and may also review your research plan to ensure the exported file(s) is/are in compliance.
- Once approved, the requested file(s) will be moved to the user's SFTP space. The researcher will then receive notification via email.

Note: Supporting documentation will not be exported but must be included in the export request for NHATS staff to review. Export requests will be rejected if they do not contain an accurate vetting request form.

A reviewer will vet the files to be exported and may also review your research plan to ensure the exported file(s) is in compliance. Reviews can take up to five business days.

NHATS Staff, in consultation with the NHATS PIs, have the final decision over whether a given set of results may be exported.

IV. LINKAGE Enclave Import Requests

Requests to bring files in or out of the LINKAGE enclave should go to the NHATS Compliance Officer (CO) using the LINKAGE enclave communications portal.

- Log in to Platforms Main.
- In the left navigation, click Communications.
- Locate the "To Project Desktop" topic for your project.
- In the Files column, click Library.
- Below the Keywords search, click the Upload Files button.
- Click the Choose File button.
- Navigate to and select files for upload (if all files are in the same folder, you can select all using CTRL + A [Windows] or CMD + A [Mac]).
- At the base of the Open modal, click the Open button.
- In the Description field, enter a brief summary of your file transfer request, including what the file or files contain and why they are needed.
- At the base of the page, click the Upload button.

The NHATS Compliance Officer will review your list of approved files to make sure the file is listed and then vet the import file(s) for content. Once approved, the requested file(s) will be moved to the requester's folder. Once the process is complete, the researcher will receive notification via email.

If the file (excluding statistical code) is not listed in your research plan and external file request form, you will be asked to update your list of approved files in a new external file request form. If the file does not directly correspond to your research plan, you will be asked to provide a revised research plan, which will have to be approved before the file can be imported.

III. LINKAGE Enclave Export Requests

Once self-vetting is complete:

- In your project's Desktop Session, open the "Analytic-Data Apps" directory.
- Open the "File Auditing and Security Tool" and click the "Project Desktop to CP" link.
- Click "New Transfer."
- Enter the reason for the transfer.
- Fill out a vetting request form as completely as possible.
- Drag and drop the vetting request form, the file(s) for review, along with supporting documentation from the project folder.
- Check attestation.
- Submit.

Note: Supporting documentation will not be exported but must be sent through FAST in order for NHATS staff to review. Export requests will be rejected if they do not contain an accurate vetting request form.

A reviewer will vet the files to be exported and may also review your research plan to ensure the exported file(s) is in compliance. Reviews can take up to five business days.

Once approved, the requested file(s) can be downloaded from the Communications Portal.

NHATS Staff, in consultation with the NHATS PIs, have the final decision over whether a given set of results may be exported.

Appendix A: Vetting Request Form (not publicly available)

Appendix B: List of Common Issues in Export Requests that Violate Vetting Rules

Rule	Why Rule Exists	Common or Potential Issues in Export Requests that Violate the Rule	Remedy for Researcher
1. Only individual level analyses are allowed	NHATS was not designed to be representative of other units of analysis	Units of analysis such as facility or state are the focus of the analysis	Eliminate all results that have a unit of analysis other than the individual
2. Table title must make clear the sample and type of results presented	Vetting Analyst/NHATS Staff need to fully understand the results requested for export	Sample / type of results are unclear	Add titles to tables
3. All rows and columns must have understandable labels	Vetting Analyst/NHATS Staff need to fully understand the results requested for export	Label unclear or coded variable names used instead of descriptions	Make row/column labels clearer
4. All unweighted table cell counts must be provided in tables requested for export or as supporting statistics	Underlying samples below 11 raised disclosure risk.	Researchers forget to provide supporting statistics	Add unweighted cell counts to table or create a table with supporting unweighted cell counts
5. All categories of a variable must be included or collapsed. NHATS does not allow cell suppression (i.e., dropping a category because it is too small)	Small numbers of missing cases that can be derived through arithmetic represent disclosure risks.	Researchers drop a variable category because it is too small	Combine rows or columns to include cases
6. Unweighted table cell counts should sum to total analytic sample size.	Vetting Analyst/NHATS Staff need to understand the analytic sample. Small changes in analytic samples can raise disclosure risk.	Variables sum to different sample sizes, indicating that there is unacknowledged missing data in certain variables	Collapse rows or columns, revise the table to reflect the final analytic sample, or impute missing cases
7. All unweighted table cell counts must be at least 11, including counts that can be obtained through arithmetic	Samples under 11 raise disclosure risk. CMS imposes this rule as well.	Unweighted cell counts are 10 or under	Collapse rows or columns or impute missing cases
8. All item missingness should be included in the table, or combined with another cell (imputed or recoded) if $n < 11$	Small numbers of missing cases represent disclosure risks.	Researchers do not note missing cases in the unweighted table and the categories do not sum to the analytic sample size	Collapse rows or columns, revise the table to reflect the final analytic sample, or impute missing cases

9. Redaction of a single cell, column, or row is not allowed	The redacted categories represented in the single cell, column, or row are known to be small by omission, raising disclosure risk	Researchers delete cells with small counts from the table and the categories do not sum to the analytic sample size	Collapse rows or columns, revise the table to reflect the final analytic sample, or impute missing cases
10. Text may not report which specific categories/values have been imputed or combined because of small sample size	The redacted categories represented in the single cell, column, or row are known to be small by omission, raising disclosure risk	Researchers identify cases that have been imputed or combined with another category in the table notes or the manuscript	Eliminate all references to cases that have imputed or combined due to small sample size
11. Specific geographic areas (e.g., state, county, tract, HRR, and any other geographic region below Census division (e.g., Middle Atlantic, Pacific, etc.) may not be reported	NHATS was not designed to make state-level comparisons or other direct geographic comparisons and we do not allow states or geographies to be identified given that and the potential disclosure risk raised.	Geographic areas below the Census division are identified in output	Eliminate all results that identify geographic areas below the Census division. Researchers can report they controlled for X policy or examined the effects of X policy. But they cannot report specific geographic areas below the Census division had X policy.
12. Output may not list or print cases. Microdata is not releasable.	The DUA specifies that only aggregate information is to be published from NHATS restricted data	Researcher wishes to export a dataset or information that they imported to the enclave	Researcher requests to export code but not microdata
13. Categorical contextual data must have at least 3 states or 11 units for lower levels of geography per category	NHATS was not designed to make state-level comparisons or other direct geographic comparisons and we do not allow states or geographies to be identified given that and the potential disclosure risk raised.	States or lower levels of geography are identifiable	Show tabulation of the contextual data showing at least 3 states per category; Eliminate all results that identify states or lower levels of geography
14. Counts (unweighted) for all variables in models must be shown	Estimates based on small cell counts are not allowed for export because they can raise disclosure risk	Unweighted model cell counts not shown in files requested for export or in supporting documentation	Add cell counts or create separate tables with cell counts for each separate analytic sample
15. In a categorical dependent variable regression, a crosstab of the predictor variable by the outcome must contain cells of 11 or more	Estimates based on small cell counts are not allowed for export because they can raise disclosure risk	Unweighted crosstabulation not shown in files requested for export or in supporting documentation	Provide tables with crosstabs for each separate analytic sample; If cells are less than 11, the researcher should collapse categories or drop variable; Researcher should not report which cell was collapsed

Appendix C: References

- Hawes, M. B., Brassell, E. M., Caruso, A., Cumings-Menon, R., Devine, J., Dorius, C., Evans, D., Haase, K., Hedrick, M. C., Holan, S. H., Hollingsworth, C. D., Jensen, E. B., Kifer, D., Krause, A., Leclerc, P., Livsey, J., Rodríguez, R. A., Rogers, L. T., Spence, M., ... Keller, S. A. (2025). ***Toward a Principled Framework for Disclosure Avoidance***. U.S. Census Bureau. <https://hdsr.mitpress.mit.edu/pub/bux70n5u>
- Lauger, A., Wisniewski, B., & McKenna, L. (2014). ***Disclosure Avoidance Techniques at the U.S. Census Bureau: Current Practices and Research***. U.S. Census Bureau.
- Statistics Canada Data Access Division Vetting Committee. (2022). ***DAD Vetting Handbook, Version 1.1***. https://crdcn.ca/app/uploads/2022/07/DAD-Vetting-Handbook_EN.pdf
- U.S. Census Bureau Disclosure Review Board. (2023). ***Federal Statistical Research Data Center Disclosure Avoidance Methods: A Handbook for Researchers, Version 4.0***. <https://www.census.gov/content/dam/Census/programs-surveys/sipp/methodology/FSRDC%20Disclosure%20Avoidance%20Methods%20Handbook%20v.4.pdf>