CCW VRDC Data Output Review Do's and Don'ts

Chronic Conditions Warehouse

Your source for national CMS Medicare and Medicaid research data

October 2023 | Version 1.1

INTRODUCTION

This document is meant as a companion to the *Chronic Conditions Warehouse (CCW) Virtual Research Data Center (VRDC) Data Output Review Process* guide that CCW VRDC users can find on the CCW website Analytic Guidance tab. It is a short list of the more common do's and don'ts CCW output review analysts address when getting users their data.

The purpose of the data output review process is to help CCW VRDC users avoid disclosure, or the perceived disclosure, of confidential information. The CCW analytical team reviews all output and ensures it meets all disclosure checks before transferring it to the user for download.

Do

Data Suppression (examples on slides 4 and 5)

- ✓ Confirm files DO NOT have patient-related counts or frequencies between 0 and 11.
  - Remember beneficiary data includes enrollment information, service counts, visits, length of stay, claims, procedures, and Part D Events.
- ✓ Check missing, censored, and death counts to make sure they are suppressed when between 0 and 11.
- ✓ Collapse categories to meet minimum sample sizes when stratifying by race.

Labeling

- ✓ Label output summary statistics, such as mean/average/ standard deviations.
- ✓ Format dollar amounts as currency (such as with a $).
- ✓ Label all headers, columns, and variables on the datasets and spreadsheets.

Don't

Data Suppression (examples on slides 4 and 5)

- X Include medians, maximums, minimums, or percentiles if the sample size is less than 50.
- X Include extreme values such as those produced by proc univariate.

Labeling

- X Vaguely label means and averages. This could cause them to be mistaken for counts and rejected.
- X Leave zip codes unlabeled.
- X Send spreadsheets or other type of datasets without header or variable labels.

File Types

- X Deviate from the accepted file types listed in the *CCW VRDC Data Output Review Process* guide.

EXAMPLES — SUPPRESS SMALL CELL SIZES

Remember to suppress small cells sizes, and any **cumulative frequencies** that allow for back calculation of small cell sizes.

Example: study of the number of individuals with dental disease and their smoking status.

- The first example shows why the CCW team will reject this file — because 8 is less than 11.
- The second example shows the data requestor has properly suppressed the count of individuals who have never smoked; however, a person can recalculate the count using data from the Cumulative frequency column (64-56 = 8).
- The CCW team will only approve the third example because by suppressing counts, it sufficiently suppresses the ability to calculate the count.

Example 1: Unsuppressed – Rejected

Smoking	Count	Cumulative frequency
current	39	39
former	17	56
never	8	64
missing	14	78

Example 2: Calculable – Rejected

Smoking	Count	Cumulative frequency
current	39	39
former	17	56
never	<11	64
missing	14	78

Example 3: Fully Suppressed – Approved

Smoking	Count	Cumulative frequency
current	39	39
former	17	56
never	*	*
missing	14	*

EXAMPLES — SUPPRESSION (CONT'D.)

Remember to suppress small cells sizes, and any **associated rates** that allow for back calculation of small cell sizes.

Example: reviews the number of individuals with dental disease and their smoking status.

- The first example shows 10 individuals were never smokers and had dental disease. The CCW team will reject this file because 10 is less than 11.
- The second example has correctly suppressed the number 10; however, a person can calculate the number of individuals with dental disease (100 x 10/100 = 10).
- The third example suppresses the number of individuals and the dental disease rate. The CCW team will only approve the third example.

Example 1: Unsuppressed — Rejected

Smoking	No. with dental disease	Count by smoking status	Rate
Current	100	400	25%
Former	50	200	25%
Never	10	100	10%
Missing	50	100	50%

Example 2: Calculable— Rejected

Smoking	No. with dental disease	Count by smoking status	Rate
Current	100	400	25%
Former	50	200	25%
Never	*	100	10%
Missing	50	100	50%

Example 3: Fully Suppressed — Approved

Smoking	No. with dental disease	Count by smoking status	Rate
Current	100	400	25%
Former	50	200	25%
Never	*	100	*
Missing	50	100	50%

SUBMIT ONLY NEEDED FILES

Do

- ✓ Complete all analysis in the CCW VRDC environment.
- ✓ Only submit *final* files — not preliminary files — for CCW VRDC output review.
- ✓ Provide only a single version of a file for submission to the File Transfer Request System (FTRS).
- ✓ Manually review every file before submitting to the FTRS.
- ✓ Ensure individual files within the request are 1 GB or less before submitting the request.
- ✓ Submit allowed geographic variables, such as county or larger for beneficiaries.
- ✓ Compress (zip) to a single folder level, if compression is necessary.

Don’t

- X Submit files that are draft or not needed outside the CCW VRDC environment.
- X Send files directly from SAS to the FTRS Output folder.
- X Resubmit the same file, or same summary, in a different file format.
- X Submit any files that are not sharable, emailable, publishable, or that pose any disclosure risk.
- X Zip folders within folders. If one file within the zip file is rejected, then the CCW team must reject the entire zip file.
- X Submit files with hidden worksheets, columns, rows, or cells.
- X Submit beneficiary zip codes, census tract, or longitude and latitude metrics; the CCW doesn’t provide census tract or waypoints.

ADHERE TO THE DATA USE AGREEMENT (DUA)

Do

- ✓ Understand the CCW DUA requirements.
- ✓ Comply with the terms of your DUA. It is a violation of the DUA to purposely manipulate files to subvert any output suppression rules.

Don't

- X Rely on the CCW output review analysts to find violations in the file.
- X Manipulate or submit files to appear as if they are not in violation.

CCW VRDC users who frequently request to download output that is not compliant with CMS's policies may have their CCW VRDC access suspended or terminated for violation of their DUA.