

Suggested Citation:

National Center for Health Statistics, Office of Analysis and Epidemiology. Use of Survey Weights for Linked Data Files – Preliminary Guidance. May 2013. Hyattsville, Maryland.

(Available at the following address:

http://www.cdc.gov/nchs/data/datalinkage/use_of_survey_weights_for_linked_data_files.pdf)

Last updated: May 2013

Use of Survey Weights for Linked Data Files – Preliminary Guidance

The survey weights provided in NCHS population health survey data files adjust for oversampling of specific subgroups and differential non-response, and are post-stratified to annual population totals for specific population domains to provide nationally representative estimates. The properties of these weights for linked data files with incomplete linkage due to ineligibility for linkage and non-matches are unknown. In addition, methods for using the survey weights for some longitudinal analyses are yet to be researched. Because this is an important and complex methodological topic, ongoing work at NCHS and elsewhere is examining the use of survey weights for linked data in multiple ways.

Until specific recommendations are available, preliminary guidance is to analyze linked data files using adjusted survey weights. The survey weights available on NCHS population health survey data files can be adjusted for incomplete linkage and non-matches, hereafter referred to as non-response, using standard weighting domains to reproduce population counts within these domains -- gender, age, and race/ethnicity subgroups. These counts are called `control totals` and are estimated from the full survey sample.

A model-based calibration approach developed within the SUDAAN software package (Procedure WTADJUST) allows auxiliary information to be used to adjust the statistical weights for non-response. This approach is promising and this software is provisionally recommended for adjusting survey weights for the linked files. As inferences may depend on the approach used to develop weights, within SUDAAN's WTADJUST or using a different calibration approach, researchers should seek assistance from a statistician for guidance on their particular project. Additional information on using Procedure WTADJUST to adjust sample weights for linkage eligibility, including sample SUDAAN code, can be found at: http://www.cdc.gov/nchs/data/datalinkage/adjusting_sample_weights_for_linkage_eligibility_using_sudaan.pdf

A detailed literature review of weighting methods will be posted on the NCHS Data Linkage webpage as it becomes available. Please contact the NCHS Data Linkage Team (datalinkage@cdc.gov) for more information.