Subsample Weights Studies - Alternate Methods for Constructing BRR Weights for NHANES Single Year Samples

Te-Ching Chen and Jennifer Parker
National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782

Abstract

Balance repeated replicate (BRR) weights are commonly used for variance estimation for multi-stage complex sample surveys. BRR requires two primary sample units (PSU) per stratum. Creating pseudo design variables, BRR stratum and/or BRR PSU, is needed when there is only one PSU per stratum. The National Health and Nutrition Examination Survey (NHANES) is a multi-stage sample survey with a sample design that has changed over time. For NHANES 2011-2014 stratification, five state groups were formed using external health information (e.g., infant mortality rates). Using these five state groups, major strata were created by geographical and urbanization characteristics. This design is used to create variance units on released data files for variance estimation; major strata and certainty PSU are both used to form variance strata. NHANES has two (sometimes three) variance PSUs per stratum for each two-year data file. For single year data files, there is generally one variance stratum and 15 variance PSU. Therefore, it is not straightforward to create BRR sample weights from single year data files. Although most health measures are available on two-year data files, some components are only available for one year. The purpose of this paper is to compare three methods for creating BRR sample weights with a one-year NHANES data file. First, two PSUs are randomly paired into one BRR stratum (Random). Second, the five state groups are used as BRR strata and PSUs within each state group are paired (State Group). Third, each PSU is defined as a BRR stratum and segments within the PSU are divided into two BRR PSU (Dividing PSU). Simulations using 16 years of NHANES data are used to examine variance estimates using these three methods. Relative bias and coefficient of variations (CV) are compared among the three proposed methods.

Key Words: BRR, Re-weighting, Variance Estimation, NHANES, pseudo-PSU.

1. Introduction

The National Health and Nutrition Examination Survey (NHANES) is a sample survey of civilian, non-institutionalized US residents. NHANES collects and releases health and nutrition data in two-year cycles, with the exception of some special studies, which only last a year or less. NHANES has a stratified multi-stage sample design, which includes primary sample units (PSU), segments, households and individuals in the households. Specific details of the design change over time, including criteria for oversampling population subgroups and forming strata. Currently, the sample design is implemented over four years, with two-year data releases (Johnson et. al. 2014). Over two-years, there are two PSUs sampled in each stratum, one each year. From this design, variance strata and variance PSUs are created for variance estimation, with two (sometimes three) variance

PSU for each variance stratum in a two-year data file. (These variance units are masked on public use files to protect the confidentiality of information.)

In a complex sampling survey setting like NHANES, variance estimates are biased if simple random sampling is assumed. Two common approaches are available for estimating variances for complex survey data with the variance units on the data files: linearization and replication (Rust and Rao, 1996). Fay's-adjusted balanced repeated replication method (Fay's BRR) is a widely used replication method. Different from other replication methods, including traditional BRR, Fay's BRR uses all observations in every replication (Judkins, 1990), which avoids instability due to extreme estimates.

BRR weights require two PSUs per stratum for implementation. It is straightforward to construct BRR weights for NHANES two-year data files. However, with only one PSU sampled in each stratum in a single year, it is necessary to construct BRR strata and BRR PSUs to create BRR weights for single year data files.

The objective of this paper is to compare alternative methods to construct BRR strata and BRR PSUs to create BRR weights for NHANES single-year data files. Simulation was used to examine the three proposed BRR weight construction methods. The construction methods were evaluated by comparing BRR variance estimates, relative biases and coefficient of variations (CV) using a variable based on obesity prevalence, a commonly reported NHANES health outcome. Data from NHANES 1999-2014 were used as the basis for the simulations. Simulations were done separately by survey year. The range and average of the simulation results across the 16 years and the simulation results for Year 2014, the most recent survey year, are shown. Additional results for Year 2014 are shown by race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic and Asian), age group (20-39 years, 40-59 years, and 60 years and over), and gender.

2. Sample Design

NHANES has a four-stage sample design with stratified PSUs from a frame of all U.S. counties (Curtin et. al. 2012, Johnson et. al. 2014). In the design, some PSUs have such a large measure of size that they are selected into the survey with a probability of one and referred to as "certainty PSUs". All other PSUs are selected without certainty and are known as "noncertainty PSUs". The first stage of the sample design consists of selecting non-certainty PSUs from nested major and minor strata formed within state-groups.

Criteria for forming strata change over time. For the NHANES 2011-2014 stratification, five state groups were formed using external health information (e.g., infant mortality rates). Using these five state groups, major strata were created by geographical and urbanization characteristics.

The second stage of NHANES sample design is the selection of segments, the neighborhood blocks, within selected PSUs. The sample was designed to have approximately equal sample sizes per PSU. Noncertainty PSUs have 24 segments, while certainty PSUs might have more or less than 24 segments to ensure appropriate representation in the sample. After the selection of segments, dwelling units (DUs) in the segments are selected. Persons within occupied DUs are selected as the fourth stage of NHANES sampling. Currently, the number of NHANES study locations each year is 15 with a target of 5000 examined participants.

Variance units for two-year data files are formed from the sampling design, where major strata and certainty PSU are used to form variance strata. There are two, sometimes three, variance PSU for each variance stratum on a two-year data file. Single year data files have one variance stratum and, typically, 15 variance PSU. The single-year data are available in the NCHS Research Data Center to reduce disclosure risks (Mirel et al, 2013).

For clarity within this paper, state group, major strata, sampling PSU, and segment indicate original sampling units. Variance strata and variance PSU refer to the units used for variance estimation on the released data files. Unless otherwise specified below, PSU and strata refer to the variance PSU and variance strata.

3. Balanced Repeated Replication (BRR)

Using replication methods, estimates for parameters of interest are calculated multiple times using variations of the original sample weights; the population variance is estimated from the set of replicated estimates. One replication method, Fay's-adjusted balanced repeated replication (Fay's BRR), is widely used. For each replicate using Fay's BRR, the weights for one PSU in the stratum are increased to 2-k times the original weights while the weights of the other PSU are decreased to k times of the original weights, with Fay's factor k, $0<k<1$. Which PSU has increased weights and which has decreased weights is based on the Hadamard matrix. In order to obtain a fully balanced design, the number of Fay's BRR replicates needs to be a multiple of four greater than the number of strata. Importantly, exactly two PSUs per stratum are required to calculate the BRR weights.

The Fay's BRR variances equals the mean square error of the estimates divided by $(1-k)^2$. The weights and variances are equal to the traditional BRR when $k=0$.

4. Methods

With two PSUs per stratum, it is easy to construct Fay's BRR weights with NHANES for the two-year samples. However, as NHANES has only one PSU per stratum in single-year samples, constructing Fay's BRR weights for these samples is not straightforward and it is necessary to create BRR strata and BRR PSU to mimic a two PSU per stratum design.

In this paper, we propose three alternative methods to construct BRR strata and BRR PSU for NHANES single-year samples: 1) Random, 2) State Groups, and 3) Dividing PSU.

- The **Random method** is the most straightforward method. It treats the PSUs as BRR PSUs and groups two of these PSUs randomly into a BRR stratum. In the case of an odd number of PSUs, the last PSU is added to one of the BRR strata and two of these three PSUs in that stratum are combined into one BRR PSU randomly with a higher probability of grouping the two smallest PSUs together than of combining the largest PSUs. The other PSU is the 2nd BRR PSU of the stratum.

- The **State Groups method** treats the five state groups as BRR strata and groups the PSUs under a state group into two BRR PSUs. Certainty PSU are assigned to state groups based on their location. The grouping technique used for the Random method is used to combine PSU within each state group.

- The **Dividing PSU method** treats the PSUs as BRR strata and groups segments within each BRR stratum randomly into two BRR PSUs.

Table 1 shows the three BRR weight construction methods used in this study.

Table 1: Units* used to create BRR strata and BRR PSU using three construction methods

	Construction Method		
	Random	State Groups	Dividing PSU
BRR Strata	Randomly grouped 2 PSUs	State Groups	PSUs
BRR PSUs	PSUs	Grouped PSUs	Grouped segments

*Sampling design units include state group stratification and sampling segments. PSU refer to variance PSU for variance estimation on released data files.

5. Simulations

Simulation was used to examine the three proposed BRR strata/PSUs construction methods. The simulation samples were generated using single year data for NHANES 1999 – 2014, assuming the NHANES 2011-2104 stratification for each year. The construction methods were evaluated by comparing BRR variance estimates, relative biases and coefficient of variations (CV) for a health variable. The health variable was created based on obesity among adults, a commonly examined NHANES health outcome. The obesity indicator was defined as $I_{obese} = 1$, if the participant's body mass index ($BMI = Kg/(m^2)$) was at least 30. For this simulation, the obesity variable was randomized across age group, race-ethnicity, and gender domains for convenience. As obesity differs by these factors, the BRR variance estimates, relative biases and CV do not correspond to national estimates for obesity among adults.

5.1 Simulation replications

The simulations were created with 16 years of NHANES released demographics and body measurement data together with restricted sampling information and restricted single-year MEC examination weights. Participants aged 20 years old and older were included in the simulation data generation frame. The single-year sample sizes ranged from 1,315 to 2,313 with a total of 31,258 participants over all years.

The data frame included three datasets: 1) personal data, 2) sample units and 3) obesity. The personal data dataset included the participant ID, survey year, demographics (gender, race/ethnicity and age group), and the single-year MEC examined weight. The sample units dataset included survey year and indicators for the NHANES sampling stages (e.g., segment), stratification scheme, variance strata and variance PSU. The obesity dataset included survey year and the derived obesity indicator, referred to hereafter as "pseudo-obesity".

Simulation replications were created separately for each survey year. To create simulation replications, the personal data dataset was sorted by survey year and participant ID and a new variable, observation order, was created within survey year. The personal data dataset remained unchanged for all 500 simulation replications. To randomize the sample-units and pseudo-obesity, two sets of random numbers were generated and attached to the sample units and obesity datasets for each simulation replication. The sample units and obesity

datasets were sorted separately by survey year and random number and new observation order numbers were created. For each simulation replication, the three datasets were merged together as the "Full" sample.

5.2 Construction of the BRR Units

5.2.1 "Random" BRR units

As mentioned in Methods section, two PSUs within the same survey year were grouped together as a BRR stratum. The PSUs were assigned as BRR PSU 1 and BRR PSU 2, alternatively, and grouped into BRR strata. When there was an odd number of PSUs in the survey year, the last three were grouped into one BRR stratum.

5.2.2 "State Groups" BRR units

The NHANES stratification scheme changes over the NHANES sample designs. The state groups used in the NHANES 2011-2014 design were applied to all of the 16 survey years to form strata in this project.

Each state group was treated as a BRR stratum for the State Groups method. In general, there were more than two PSUs within a state group. To combine the PSU into two groups, a technique similar to that used to assign BRR PSUS for the Random method was used for State Groups method. All of the PSUs assigned to BRR PSU 1 in a BRR stratum (state group) were grouped together as one BRR PSU and all of the other PSUs in the same BRR stratum (state group) were grouped together as the other BRR PSU.

5.2.3 *"Dividing PSU" BRR units*

The PSUs were used as BRR PSUs in the Random method and the State Groups method but were used as BRR Strata for the Dividing PSU method. The second stage sampling units, segments, were used to form BRR PSUs for the Dividing PSU method. To implement this method, a set of random number was attached to the list of segments within survey years and BRR Strata (PSUs). For each survey year and BRR Stratum, the segment list was sorted by the random number. Segments were assigned into one of two BRR PSUs within the BRR Stratum alternatively.

6. Simulation Analysis

The BRR units for the three approaches (Random, State Group, and Dividing PSU) were merged back to the original data frame by survey year and sampling units. The BRR weights were initially created using the SAS 9.3 SURVEYMEANS procedure with Fay's factor = 0.3. The initial weights then were post-stratified to the population totals of the gender-race/ethnicity-age group for each survey year.

SUDAAN SAS Callable 11 DESCRIPT procedure was used to estimate the single-year BRR variances the health variable (pseudo-obesity) using the three methods. Single-year results were obtained overall and by gender, race/ethnicity and age group. We compared the simulation results in terms of relative biases and coefficient of variations. The overall results by survey year and the results for Year 2014, the most recent year of NHANES data, by demographic subgroup are shown in this paper.

The variance estimated for pseudo-obesity from the 500 simulation samples was defined as the population variance and is referred as the "Reference" variance below.

For each of the three BRR methods, the average BRR variance was calculated as the mean of the BRR variances of the 500 simulation replications. The standard deviation of the BRR variance was calculated as the standard deviation of the BRR variances of the 500 simulation replications.

For each BRR method, the Relative Bias was calculated as the ratio of the average bias, the difference between the average BRR variance and the "Reference" variance, to the "Reference" variance (Chen et al. 2007). The coefficient of variation was calculated as the ratio of the standard deviation of BRR variance to the average of the BRR variance.

7. Findings

The "Reference variance", estimated from the 500 simulation samples, for pseudo-obesity ranged from 0.4 to 1.3 over the 16 survey years.

Figure 1 shows the overall relative biases for Survey Year 2014 and the minimums, maximums and averages of the 16 relative biases for each year for the three BRR methods. The relative biases were similar, ranging from 0.8 to 2.6, among the three BRR methods. As the relative bias is defined relative to the "Reference" variance, the minimum relative bias of 0.8 obtained using the Dividing PSUs method means that even the BRR variance estimate closest to the "Reference" value was almost twice the "Reference" variance. The average BRR variances for all methods was almost 1.5 times higher than the "Reference" variance.

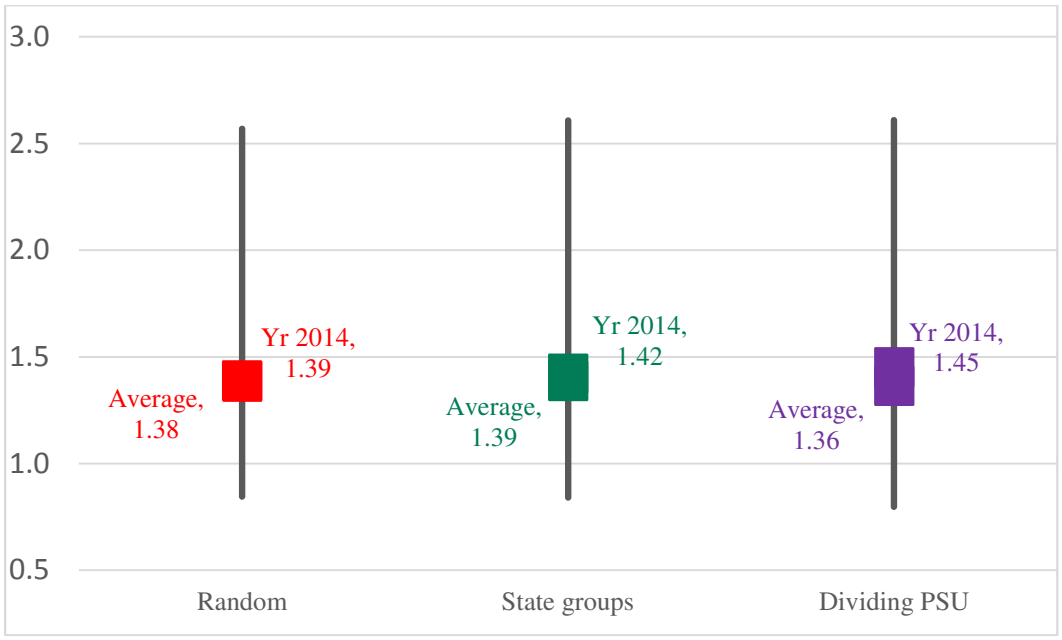

Figure 1: Minimums, maximums, averages of relative biases for pseudo-obesity variable over 16 single-year estimates and relative bias for Survey Year 2014, by BRR weight method.

The biases of the BRR variances using the three BRR weight methods were smaller by race/ethnicity, gender and age group, in part because the "Reference" variances for these subgroups were larger than the overall single-year variances (Figure 2). For example, for Survey Year 2014, the "Reference" variances by race/ethnicity ranged from 1.9 (non-Hispanic white) to 7.7 (Asian). Using the three BRR weight construction methods, variances by race/ethnicity ranged from 2.6 (non-Hispanic white, State groups method) to 8.6 (Asian, State groups method). Differences between the "Reference" variances and those calculated using the three BRR methods were smaller by gender; for Survey Year 2014, the "Reference" variances were 2.0 and 2.5 for female and male, respectively, and the variances using the three BRR weight construction methods ranged from 2.7 (female, Random method) to 3.2 (male, Dividing PSU method).

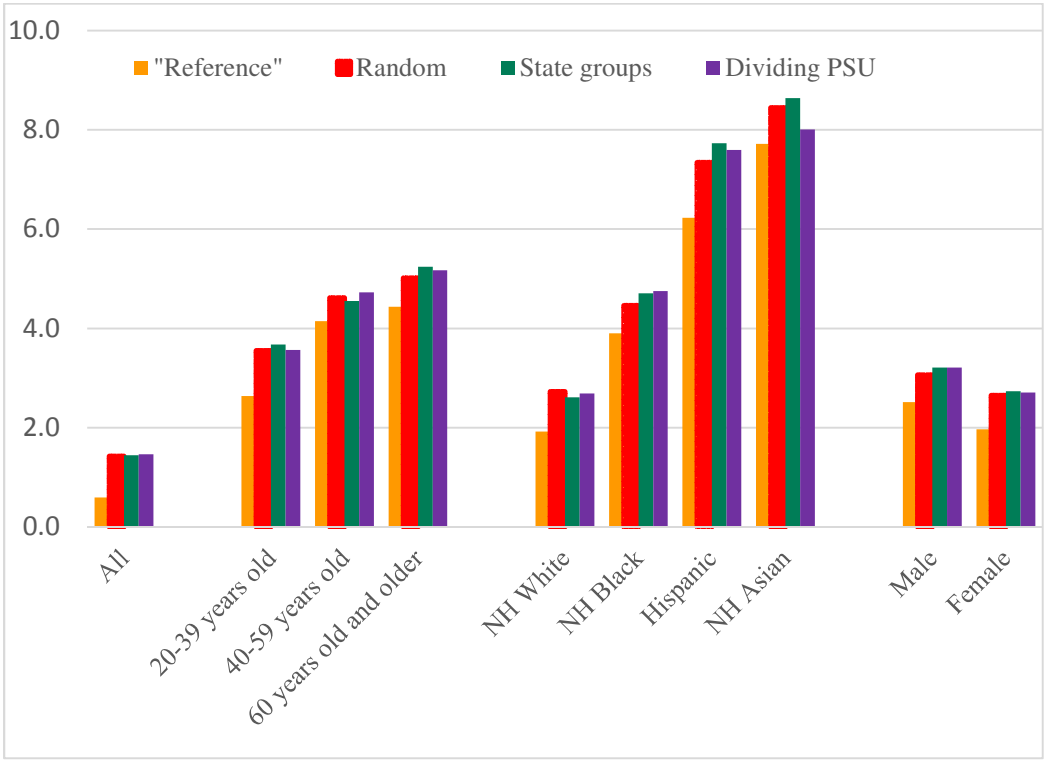

Figure 2: Variance estimates for pseudo-obesity variable by BRR weight method using one year of data (Survey Year 2014): overall and by age group, race/ethnicity, and gender.

Figure 3 summarizes the 16 single-year coefficient of variations (CV) from the three BRR weight construction methods. The CVs using the Dividing PSU method were the smallest among the three construction methods. Using the Dividing PSU method, the range of CVs for the 16 survey years was from 0.36 to 0.43 with an average of 0.38. The CVs ranged from 0.49 to 0.59 and from 0.62 to 0.79 using the Random method and the State Groups method, respectively.

CVs for gender, race/ethnicity and age groups had similar patterns as those overall when using the same method; see results for Survey Year 2014 in Figure 4.

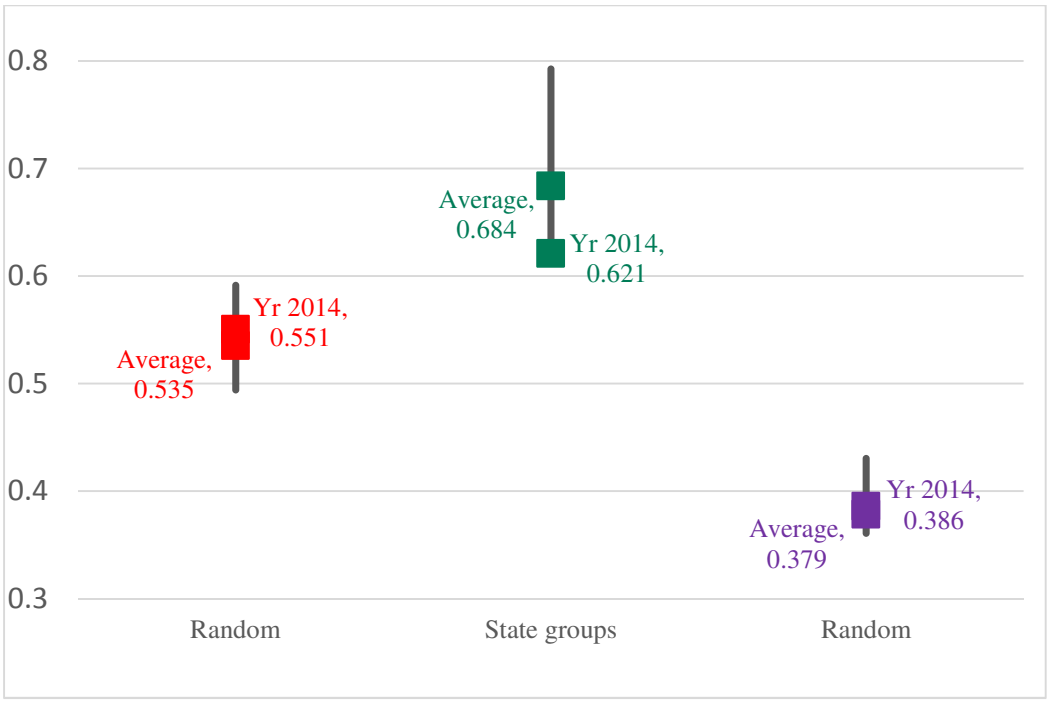

Figure 3: Minimums, maximums, averages of CV for pseudo-obesity variable over 16 single-year estimates and for Survey Year 2014, by BRR weight method.

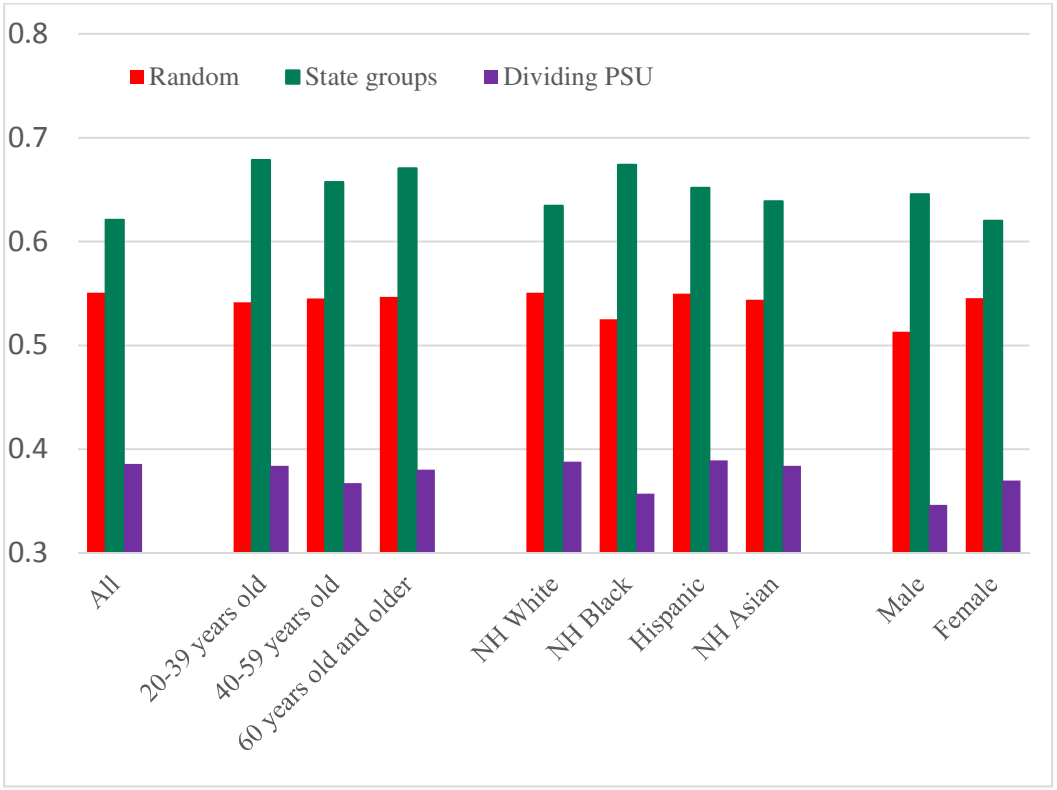

Figure 4: Variance CV for created obesity variable by BRR weight method using one year of data (survey year 2014): overall and by age group, race/ethnicity, and gender.

8. Discussion and Conclusion

There are additional features of each method to note here. The **random method** is the most straightforward method, does not require sampling information (e.g. state-group or segment) other than the variance units on the data file and is the easiest method to use to construct BRR units. The **state group method** has the least number of strata and, from the simulation results, had the largest CVs. The **dividing PSUs method** had the smallest CVs from the simulation results but might underestimate the within PSU variation as each BRR PSU contains half the number of segments in the PSU.

These simulation results were based on assuming full single-year survey samples and findings could differ for components measured on a population subgroup or for part of a survey year. In addition, these initial results were based on a health outcome without consideration of known differences in the outcome by age group, race/ethnicity, and gender. Findings could differ when the relationships between outcomes and subgroups of interest are considered.

The simulation results showed that the relative biases of the variance estimates calculated using the three methods for BRR weight construction were similar, but large. The relative biases were smaller by domains, in part due to larger "Reference" variances within the domain subgroups compared to the national sample. Although the CVs using the Dividing PSU method were much smaller than those obtained using the other two construction methods, all of the CVs were greater than 35%.

NHANES single year data files are only available in the NCHS Research Data Center. Estimates from these files are known to have large variances when using linearization methods (Mirel el at, 2013). The simulation results from this study confirmed that the variances estimates for single-year NHANES data can be big, regardless of the method used to obtain variances.

References

Chen, T.-C., P. A. Bobbitt, J. A. Himelein Jr., et al. 2007. "Variance Estimations for International Price Program Indexes", The Joint Statistics Meeting Proceedings (2007).

Curtin L. R., L. Mohadjer, S. Dohrmann et al (2012). "The national Health and Nutrition Examination Survey: Sample design, 1999-2006". National Center for Health Statistics. Vital Health Stat 2(155). Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_155.pdf

Johnson C.L., S. M. Dohrmann, V. L. Burt and L. K. Mohadjer. 2014. "National Health and Nutrition Examination Survey: Sample design, 2011-2014". National Center for Health Statistics. Vital Health Stat 2(162). Available from: http://www.cdc.gov/Nchs/Data/Series/Sr_02/Sr02_162.Pdf

Judkins, D. R. 1990. "Fay's Method for Variance Estimation", Journal of Official Statistics, Vol. 6, No. 3,223-239

Mirel L.B., L. K. Mohadjer, S. M. Dohrmann, et al. 2013. "National Health and Nutrition Examination Survey: Estimation procedures, 2007-2010". National Center for Health Statistics. Vital Health Stat 2 (159). Available from: http://www.cdc.gov/nchs/data/series/sr_02/sr02_159.pdf

Rust K.F. and JNK Rao (1996). "Variance estimation for complex surveys using replication techniques", Statistical Methods in Medical Research, 5, 283-310.