A Comparison of Methods to further adjust for non-response due to missing lab data in National Health and Nutrition Examination Survey (NHANES)

Maya Sternberg[1], Alula Hadgu[2]

[1]Division of Sexually Transmitted Diseases, Centers for Disease Control and Prevention, 1600 Clifton Rd. MSE63, Atlanta, GA, 30333

[2]Division of Sexually Transmitted Diseases, Centers for Disease Control and Prevention, 1600 Clifton Rd. MSE63, Atlanta, GA, 30333

Abstract

The National Health and Nutrition Examination Survey (NHANES) is designed to provide unbiased estimates of the number and percent of persons in the United States population with selected diseases and risk factors. This paper investigates the utility of additional non-response adjustments to the standard NHANES weights in order to estimate the prevalence of 6 sexually transmitted diseases (STDs).

KEY WORDS: NHANES, nonreponse, propensity

1. Introduction

NHANES is a probability sample based on a complex, multistage survey of a nationally representative sample of the US civilian noninstitionalized population. Some populations, such as adolescents, African-Americans, and Mexican Americans, are over-sampled by design. Persons selected for the survey are interviewed and undergo a health examination in the mobile examination centers (MEC) (Ezzati et al). In 1999, NHANES was redesigned to become a continuous survey without a break between cycles; the details of which have been previously published (Botman et al., Montaquila et al).

Since the 1999 redesign, three two-year cycles have been publicly released and are available to download from www.cdc.gov/nchs/about/major/nhanes/datalink.htm. With each cycle, NHANES provides a standard 2-year interview and examination weight. These weights account for over-sampling, under-coverage and non-response to the interview and health examination components of the survey (Botman et al, Mohadejer et al., Ezzati and Khare).

The NHANES group at the National Center for Health Statistics (NCHS) does perform nonresponse adjustments to the original base weight. The base weight is the inverse of the probability of being selected, and adjustments made to the base weight to develop the 2-year interview and examination weight include nonresponse at both the interview and examination levels of the survey.

However, neither the interview nor the examination weights take into account non-response to, or missing, individual components of the health examination collected through the laboratory component. We investigated the impact of unavailable results on prevalence estimates of each of six sexually transmitted diseases by comparing the results of weighting class adjustments, weighted and unweighted logistic propensity models, and further post-stratification adjustments. We then compared the prevalence estimates based on the standard examination weights, to the prevalence estimates based on several different types of nonresponse adjustments to this weight.

1.1 Sexually Transmitted Diseases in NHANES

Lab findings for six sexually transmitted diseases: Herpes Simplex Virus II (HSV2), *T. vaginalis*, bacterial vaginosis, human papilloma virus (HPV), chlamydia and gonorrhea were collected on different sub-groups of participants, and through various specimen collection mechanisms, and varied cycles depending on the STD. HSV2 serology results were collected among all participants aged 14-49 who were examined during NHANES 1999-2004; self-collected vaginal swabs were tested for *T. vaginalis* and bacterial vaginosis among all female participants aged 14-49 who were examined during NHANES 2001-2004; self-collected vaginal swabs were tested for over 30 types of HPV among all female participants aged 14-59 who were examined during NHANES 2003-2004; and urine specimens were tested for chlamydia and gonorrhea among all participants aged 14-39 examined during NHANES 1999-2002. We found that 8.6%, 16.1%, 16.%, and 19.5% did not have HSV2, *T. vaginalis*, bacterial vaginosis, and HPV results, respectively, and that 3.7% did not have results for chlamydia and gonorrhea. The reasons that results were unavailable included refusal by that person to have specimen taken, unsuccessful venipuncture or specimen collection, the need to use the specimen for other tests, unusable specimens, and the loss of specimen samples during processing. For the focus of this paper we present the results from NHANES 1999-2002 for HSV2. However, the conclusions will be generalized to the other

STDs. For a complete description of HSV2 methods and results in NHANES see Xu et al.

2. Nonresponse Adjustment Methods

Two general methods are considered: weighting class adjustments (WCA) and propensity models using logistic regression. In both methods, the examination weight provided by NHANES will be further adjusted by multiplying the appropriate adjustment factor against the examination weight. We focus on adjusting the examination weight since the laboratory specimens are collected at the health examination phase and the examination weight has already been adjusted for nonresponse at the interview level. The adjustment factor increases the examination weights for the responders so they account for the non-responders. For example, suppose the probability of having a HSV2 specimen result is a function of only race/ethnicity. If 95% of non-Hispanic whites have a specimen, then each non-Hispanic white respondent with a HSV2 specimen represents $1/0.95 = 1.05$ observations. One could also compute the adjustment factor based on the inverse of weighted response rate.

2.1 Weighting Class Adjustments

To apply this method one identifies several "class" variables which are available for both responders and nonresponders. These variables are cross-classified to create cells which contain individuals who are relatively homogeneous with respect to the probability of responding to the survey. Then an adjustment factor is calculated for each cell and multiplied against the base weight of individuals in that cell. The adjustment factor is the inverse of the response rate for each cell (Korn and Graubard).

2.2.1 Weighting Class Adjustments applied to NHANES

We formed 40 weighting class cells by cross-classifying age, race, and sex and used the inverse of the weighted response rate. Only 14-59 year olds were considered, since lab STD information is not collected on people outside these age groups. None of the cells contained fewer than 20 people, so no additional modifications were made to the cells to maintain a minimum cell size. For subsequent discussion, this adjustment will be referred to as WCA.

2.2 Propensity Models

Another approach to perform nonresponse adjustment is to develop a logistic regression model that predicts the likelihood of response (Rosenbaum and Rubin, Korn and Graubard). Variables available for both responders and non-responders are used to develop a logistic regression model, which in turn is used to obtain predicted probability to respond for each individual. The inverse of this predicted probability is then multiplied by the base weight to obtain the nonresponse adjusted weight. Other propensity based strategies, not considered in this paper, are: propensity weighting and propensity stratification, which involve grouping the predicted probabilities into a number of quantiles or classes.

Whether to use a weighted or unweighted logistic regression to develop the propensity model or use the inverse of the weighted response rate for WCA has not been thoroughly discussed in the literature. An argument can be made that weights are not necessary since the goal of the propensity model is identify responders and non-responders who are similar with respect to the variables in the logistic regression model and not to make inferences about population level characteristics. On the other hand, when using statistical inference to determine which variables should or should not be included in the propensity model both the weights and the design information are relevant. For example, weights may contain information about covariates that are unavailable to public users and the design information is directly related to the appropriate estimation of the standard error and total denominator degrees of freedom for the purpose of statistical inference. Therefore, one could also consider using the design information and weights in the actual propensity model as predictors.

2.2.1 Propensity Methodology as Applied to NHANES

We reviewed the bivariate associations for 17 different variables in NHANES, using the examination weights and relevant design information for the variance calculation and inference. The variables considered included: age group, race/ethnicity, ever had sex, poverty income ratio, body mass index, recode of the strata and PSU variables to identify high/low response areas, gender, marital status, education status , current health insurance, current household food security, self-reported health status, routine place for healthcare, how often seek healthcare, type of housing, birthplace. For categorical variables with item non-response, an additional category was added to guarantee that all the examination weights for patients who received a medical examination would receive a nonresponse adjustment factor.

All variables with a Wald F p-value for the bivariate association less than or equal to 0.10 were simultaneously entered into a weighted logistic model in SUDAAN version 9.0, and any variables with a Satterthwaite Adjusted F p-value > 0.05 were dropped. This step was

followed by all including all pairwise interactions of the remaining significant variables, one at a time, and if the interaction had a Satterthwaite adjusted F <=0.05 it was retained. Once the final propensity model was identified, we reviewed the resulting distribution of propensity scores between the nonrespondents and respondents to confirm the existence of sufficient overlap between the distributions, as well as confirmed that within quartiles of the propensity score, the distribution between responders and nonresponders of each of the covariates in the final model were similar.

To obtain the nonresponse weight adjustment, the inverse of the resulting propensity was multiplied against the examination weight. Post-Stratification was also performed to the adjusted weights by ratio adjusting the weights by age, race and sex to external estimates provided by the U.S. Census Bureau. Since the examination NHANES weights are already post-stratified to these domains, the ratio adjustment factors were calculated from the examination weights summed over the appropriate domains.

2.2.2 Propensity Models used for nonresponse adjustment

The propensity model derived from the weighted logistic regression modelling steps described above included the following variables: age group, race/ethnicity, ever had sex, poverty income ratio, body mass index, recode of the strata and PSU variables and a race by strata/PSU recode

interaction. For subsequent discussion this model will be referred to as Model 1. We also considered an adjustment based on Model 1's results plus a post-stratification step, referred to as Model 1 + Post. After reviewing the distribution of the propensity scores between the nonrespondents and respondents, we determined that a model which excluded body mass index improved the overlap of the propensity distributions in the tails of the distribution. This model will be referred to as Model 2. Lastly, we derive propensity scores by fitting an unweighted logistic regression to with the same variables as Model 1.

3. Results

We present the summary of the results for HSV2. The impact of the nonresponse adjusted weights was considered by assessing the relative percent change of the estimates and the ratio of the standard errors over various variables in NHANES. In addition, each prevalence estimate that was estimated with the adjusted weight was compared to the 95% confidence interval of the prevalence estimated using the original NHANES examination weight, and if the adjusted weighted prevalence fell outside this 95% confidence interval, it was flagged.

Figure 1 and 2 show the distribution of the propensity scores for the respondents and nonrespondents for Model 1 and Model 2.

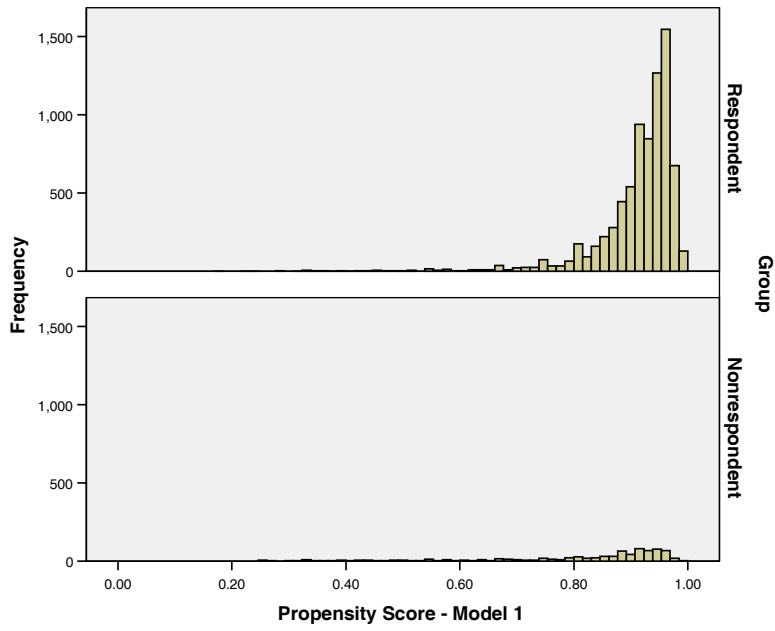**Figure 1: Propensity scores from Model 1**

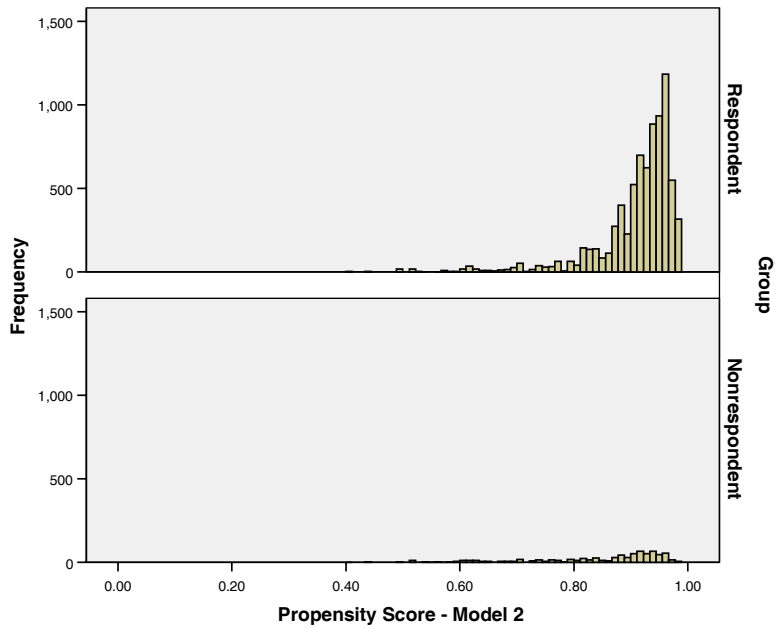**Figure 2: Propensity scores from Model 2**

Table 1 provides the unweighted and weighted point estimates for the prevalence of HSV2 by race/ethnicity, as well as the 95% confidence intervals for the weighted prevalence. The confidence intervals for the prevalence estimates were calculated using a log transformation with the standard error of the log prevalence based on the delta method (Korn and Gaubard) and applying the SUDAAN estimated standard errors based on Taylor series linearization. With the exception of the "other race" category, the unweighted prevalence estimates do not fall within the 95% confidence intervals, demonstrating the relevance of using the appropriate weights and design information when estimating population prevalence of HSV2. Most notable is the difference in the unweighted and weighted prevalence of HSV2 among non-Hispanic blacks, 27.5 verses 41.1% respectively. The weighted estimate is 50% larger than the unweighted estimate, suggesting the examination weights for those who tested positive for HSV2 tended to be higher than the average examination weight among non-Hispanic blacks.

Table 2 shows the weighted prevalence estimates and 95% confidence intervals of HSV2 by race/ethnicity after applying the four nonresponse adjustments. It is clear that the impact of change on the prevalence estimate is minimal with change only in the first decimal place of the percent. The relative percent change from the original weighted prevalence estimate across the race/ethnicity categories ranges from -1.2% to 1.1%. In addition, the width of the confidence intervals remains almost identical to those shown in Table 1. In addition, the added post-stratification step does not appear to have much of an effect, nor does dropping a significant predictor, body mass index, from the logistic model.

We considered several variables, such as: age, race, sex, number of lifetime partners, general health, education, place of origin, and calculated the adjusted weighted prevalence. All the adjusted weighted estimates fell inside the 95% confidence intervals of the original examination weighted estimate. Table 3 shows the mean relative percent change and the range of the adjustment. These values are summarized by averaging over categories of selected variables. The estimated standard errors for the adjusted weighted prevalence was almost always the same as the standard errors with the original examination weights and the maximum relative percent change was no more than 5% of the original examination weighted estimate.

Table 1 Unweighted and Weighted HSV2 prevalence estimates using original examination weight

Race/ethnicity	HSV2 Unweighted Prevalence (%)	HSV2 Weighted Prevalence (%)	95% CI using log transform for weighted prevalence
Overall	14.8	17.7	(15.9, 19.8)
NH-White	11.3	14.3	(12.7, 16.0)
NH-Black	27.5	41.1	(37.3, 45.3)
Mexican-American	9.2	12.8	(10.8, 15.1)
Other	15.6	17.7	(13.4, 23.3)

Table 2 Weighted HSV2 prevalence estimatesusing laboratory nonresponse adjusted examination weights

	Model 1		Model 1 + Post		Model 2		Unweighted Model 1		WCA	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI	%	95% CI
Overall	17.8	(16.0, 19.8)	17.8	(16.0, 19.9)	17.8	(15.9, 19.8)	17.7	(15.9, 19.7)	17.8	(15.9, 19.8)
NH-White	14.2	(12.7, 16.0)	14.2	(12.7, 15.9)	14.1	(12.6, 15.9)	14.2	(12.6, 15.9)	14.1	(12.5, 15.9)
NH-Black	41.0	(37.1, 45.2)	41.6	(37.7, 45.9)	41.0	(37.1, 45.2)	41.0	(37.2, 45.2)	41.1	(37.3, 45.4)
Mex-Am	12.6	(10.7, 14.9)	12.7	(10.7, 15.0)	12.6	(10.7, 14.9)	12.7	(10.7, 15.0)	12.7	(10.7, 15.0)
Other	17.5	(13.3, 22.9)	17.3	(13.2, 22.7)	17.6	(13.4, 23.1)	17.6	(13.4, 23.1)	17.6	(13.4, 23.1)

Table 3 Summary of relative percent change and ratio of the standard error across various categories of NHANES variables

Adjustment Method	Mean Adj SE / Orig SE	Range Min - Max	Mean Rel Δ= 100 x (Adj-Orig)/Orig (%)	Range Min - Max
Model 1	1	0.94 - 1.06	0.5	-3.22 - 4.18
Model 1 + Post	1	0.95 - 1.06	0.4	-2.92 - 3.54
Model 2	1	0.94 - 1.04	0.2	-3.78 - 2.77
Unweighted Model 1	1	0.98 - 1.02	0	-2.27 - 1.70
WCA	1	0.97 -1.01	0	-2.24 - 1.52

4. Conclusions

As demonstrated in the results for HSV2, the additional nonresponse adjustment has little practical effect on the estimated prevalence estimate and, for the variables considered, always fell within the original 95% confidence interval of the examination weighted estimated. Therefore, we conclude that no further non-response adjustment is needed to the examination weight to account for the laboratory nonresponse for sexually transmitted diseases. These results generalized to investigating the nonresponse among the other sexually transmitted disease.

Some limitations of this analysis include: the possible exclusion of a variable that may be related to nonresponse that is either unavailable or has too much item nonresponse to make the variable helpful. For example, public users do not have access to the originally selected sample and any of the household data collected from the nonrepondents to the interview. In addition, important variables such as the actual design information, region of the country, and whether the area is a metropolitan area are currently not available in the publicly released data. In addition, both approaches rely on the untestable assumption that the nonresponse is missing at random (Little and Rubin), though this assumption may be reasonable for laboratory data. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

References

Botman, S.L., Moore, T.F., Moriarity, C.L., Parsons, V.L. (2000). Design and Estimation for the National Health Interview Survey, 1995-2004. National Center for Health Statistics. Vital Health Statistics. 2(130).

Ezzati, T. and Khare, M. (1992). "Nonresponse adjustments in a National Health Survey," 1992 *Proceedings of the American Statistical Association*, Survey Methods Section, Alexandria,VA: American Statistical Association: 339-344.

Ezzati, T., Massey, J. Wakesberg, J., Chu, A., Maurer K. (1992) Sample Design: Third National Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Statistics 2(113).

Korn E., Graubard, B. (1999). Analysis of Health Surveys. New York: Wiley.

Little, RJ and Rubin DB. (1987) Statistical Analysis with Missing Data. John Wiley and sons: New York.

Mohadejer, L., Montaquila, J., Wakesberg, J., Bell, B., James, P., Flores-Cervantes, I., Montes, M (1996). National Health and Nutrition Examination Survey III. Weighting and Estimation Methodology Executive Summary.

Montaquila, J., Mohadejer, L., Khare, M. (1998). "The Enhanced Sample Design of the Future National Health and Nutrition Examination Survey," 1998 *Proceedings of the American Statistical Association*, Survey Methods Section, Alexandria, VA: American Statistical Association: 662-667.

Research Triangle Institute (2004). SUDAAN Language Manual, Release 9.0, Research Triangle Park, NC. Research Triangle Institute

Rosenbaum, P.R and Rubin, D.B. (1983) Reducing bias in observational studies using subclassification on the

propensity score. Journal of the American Statistical Association, 79: 516-524.

Xu, F., Sternberg, MR, Kottri, B., McQuillan GM., Lee, FK, Nahmias, AJ., GM., Berman, SM, Markowitz , LE. (2006). Trends in Herpes Simplex Virus Type 1 and Type 2 Seroprevalence in the United States. *JAMA*. 296: 964-973.