MODELING OF RESPONSE PROPENSITY IN THE THIRD NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY

Trena M. Ezzati-Rice, Meena Khare, National Center for Health Statistics
Trena M. Ezzati-Rice, 6525 Belcrest Road, Hyattsville, MD 20782

Key Words: Nonresponse adjustments, Weighting, Logistic regression

Introduction

The National Health and Nutrition Examination Survey (NHANES) is a periodic national survey conducted by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention. The NHANES is designed to provide national statistics on the health and nutritional status of the civilian noninstitutionalized population through household interviews and standardized physical examinations. The physical measurements and physiological tests are conducted in specially equipped mobile examination centers (MECs) that are transported to each survey location. The Third National Health and Nutrition Examination Survey (NHANES III) is the seventh in a series of surveys using health examination procedures that have been conducted since 1960 by NCHS. The target population for the survey is the civilian noninstitutionalized population aged 2 months and older. The NHANES III with a sample of approximately 40,000 persons has been divided into two 3-year surveys (phase 1 and phase 2) so that national estimates can be produced for each 3-year period as well as for the total 6 years (1988-94).

The NHANES III survey is based on a complex, multistage area probability sample design. Children under 5 years of age, older Americans aged 60 years and over, Mexican Americans, and African Americans are sampled at a higher rate than other persons. Details of the NHANES III multistage sample design, survey components, and data collection procedures have been previously published (1-2).

The NHANES, like most sample surveys, experiences unit or total nonresponse despite special procedures designed to maximize response rates. These procedures include extensive publicity in each survey location, a home examination especially targeted for the older population, a remuneration to all examined respondents, and a report of major medical findings to each survey participant. Since NHANES includes both an interview and an examination component, two levels of unit nonresponse occur--that is, some persons randomly selected for the survey refuse to be interviewed and/or examined. NHANES III-phase 1, conducted from October 1988 to October 1991, included 20,277 sample persons. The overall nonresponse rates for the interview and physical examination were 14% and 22%, respectively. For adults 17 years and older, the interview and examination nonresponse rates were slightly higher at 18% and 27%, respectively.

It is common practice to attempt to compensate for total (unit) nonresponse in sample surveys by some form of weighting adjustment. There are several weighting adjustment methods which may be used (3-6). Weighting class adjustments are commonly used in large-scale government sponsored sample surveys. This method assumes that sample persons can be divided into homogenous cells (weighting classes) within which the responses of nonrespondents, if obtained, would have been similar to those of the respondents. Within each weighting class cell, the inverse of the response rate is used to adjust the basic sampling weights for the respondents within that cell. The variables used to create the weighting classes must be available for both respondents and nonrespondents. However, to ensure adequate cell sizes, only a limited number of variables can be used to form the weighting class cells. A two-stage nonresponse adjustment for NHANES III, phase 1, using sample weighting class adjustment methodology, has been previously described (7-8). An alternative approach for forming appropriate adjustment cells and adjusting for nonresponse is to use logistic regression to model response propensity (5-6, 9-13). A response indicator variable is regressed on potential covariates that are available for both respondents and nonrespondents to get predicted probabilities. Then, as discussed in a previous paper by Little, propensity stratification or the direct use of the predicted probabilities can be used to perform the weight adjustment (6).

The purpose of this paper is to describe an evaluation of three alternative logistic regression models to predict response probabilities among adults 17 years and older for the physical examination component of NHANES III, phase 1, based on information collected from the household screener and personal household interviews. Weighted estimates for selected survey components based on response probability weight adjustments and weighting class nonresponse adjustments are then compared.

Methods

For NHANES III, two separate sets of analysis weights are developed: 1) a nonresponse adjusted interview weight for use when analyzing the household interview data, and 2) a nonresponse adjusted examination weight for use when analyzing the physical examination data. In a previous paper, we discuss a two stage weighting class model to calculate interview and examination weights for NHANES III, phase 1. SI-CHAID (Statistical Innovation's Chi-Square Automatic Interaction Detection) and logistic regression were used to determine variables related to examination response (8). For the first stage adjustment, age, race/ethnicity, geographic region, SMSA status of the survey location, and household size were used to adjust for interview nonresponse. Then, at the second stage, family income and self-reported health status were used in addition to age, race/ethnicity, and household size to adjust for examination nonresponse among interviewed persons. At this stage, region and SMSA status were excluded from the model due to the limitation on the number of weighting classes and the cell sizes.

In this paper, we use logistic regression to model response propensity and to form nonresponse adjustment cells. First, bivariate associations between examination response, and selected socio-demographic and health variables were investigated to identify predictors of response. Then, logistic regression models using a combination of potential predictor variables were compared for the goodness-of-fit. Multiple correlation coefficients (R^2) and likelihood ratio test statistics were used to test the goodness of fit of each model. Three logistic regression models were selected for evaluation. Model 1 included basic demographic variables -- age, gender, race/ethnicity, geographic region and SMSA status and household size. Model 2 included the same variables as the weighting class adjustment procedure--age, race/ethnicity, household size, family income, and self-reported health status. For this model, gender was excluded and age was grouped into two categories (<60 years and $>= 60$). The final model (Model 3) included age, age^2, gender, age*gender, race/ethnicity, geographic region, SMSA status, household size, family income, a flag for imputed income, a flag for missing education and marital status, self-reported health status, and a health index variable combining several health conditions related to nonresponse. These models identified the set of known factors most related to examination response among interviewed adults 17 years and older in NHANES III, phase 1.

As for the NHANES III weighting class procedure, two successive stages of response propensity modeling were implemented. The first stage used only demographic data from the household screener and weighted up interview respondents to all sample persons. Then, at the second stage, examination respondents were weighted up to all interviewed respondents using demographic, socio-economic and medical history information. In contrast to the weighting class adjustment method in which only a few variables can be used due to the restriction on the number of cells and cell sizes, more variables can be included in the response propensity models. Furthermore, continuous variables, like age, can be used in the propensity models as compared to only categorical variables in the weighting class procedure. The focus of this paper is on the second stage adjustment, i.e., adjustment for examination nonresponse among interviewed sample persons. Response probabilities, p(x) were estimated from each model (models 1, 2, and 3 as described earlier) from the regression of the response indicator variable (1, if the sample person was examined and 0, if not examined) on the predictor variables. The sample was then stratified by the response propensity to form adjustment cells. With the heavy oversampling in NHANES III by race/ethnicity, the propensity stratification was done within three race/ethnicity groups--Black, Mexican-American, and White and all other. This allowed us to better ensure monotone response rates within the propensity strata. The 10,120 interviewed sample persons were grouped into 15 cells with each cell containing about 750 persons. For model 2--among Whites and all others, the response rates in the cells ranged from 94% to 68%; for Blacks from 97% to 82%; and for Mexican-Americans from 94% to 87%. The adjustment factor in a cell, as for the weighting class method, was the inverse of the response rate. The nonresponse adjustment factors for Whites and all others ranged from 1.06 to 1.47; from 1.03 to 1.22 for Blacks; and from 1.06 to 1.14 for Mexican-Americans. The adjustment factors were then multiplied by the basic weight for each interviewed sample person within a given cell to produce a nonresponse adjusted weight. The final analysis weight for each sample person reflected the four-stage hierarchical sample design of NHANES III, the adjustment for nonresponse, and a final ratio adjustment to independent control estimates by age, gender, and race/ethnicity based on 1990 Current Population Survey estimates from the U.S. Bureau of the Census.

Results

Multiple correlation coefficients (R^2) and likelihood ratio test statistics were used to test the goodness of fit of the three models evaluated. Model 3 had the largest

R^2 (R^2 = 0.998). The low R^2 values observed are probably due to the high overall examination response rate among interviewed adults (87%). Further, the low values, in addition to indicating that the models do not include good predictors of response, probably indicate that examination nonresponse occurs at random among interviewed persons.

Table 1 shows weighted estimates for seven selected components from the NHANES III examination using the three logistic regression models described in the methods section to calculate predicted response probabilities and to form the nonresponse adjustment cells. The seven measurements reflect different components of the physical examination and include: two anthropometric measurements--height and weight; three blood specimen measurements--hemoglobin, serum iron, and total serum cholesterol; total bone mineral density, a measure of osteoporosis of the hip; and a macular degeneration score, a measure of visual impairment. For these seven measurements, there were no significant differences in the point estimates or standard errors among the three response propensity models used. Since response propensity model 2 included the same variables as for the weighting class adjustment model, estimates from this model were compared to estimates based on the weighting class adjustment. As shown in table 2, the two sets of weighted estimates based on the weighting class method and response propensity approach were very similar. For each of the seven selected survey measurements, the percent difference was less than 1%. Table 3 shows the distribution of the final examination weights from the weighting class and response propensity (model 2) methods. Since oversampling among selected minority subgroups is done in NHANES III, the distribution of these weights was examined within each race/ethnic category. For all three race/ethnic groups, the minimum values from the two methods were similar. However, the maximum values were much higher for the response propensity method. This is due to the fact that extreme nonresponse adjusted weights were trimmed for the weighting class method and not for the propensity method. Consequently, the CVs are also somewhat higher for the propensity method. The minimum, maximum, and mean overall adjustment factors (nonresponse and poststratification) applied to the basic sampling weight for both the response propensity and weighting class procedures are shown in table 4. The overall weighting class adjustment factors ranged from 0.34 to 2.62, while the factors from the response propensity approach ranged from 0.92 to 2.24. The mean adjustment factor within each race/ethnic group was nearly the same for each approach. We also examined the variances of the estimates for the two weighting adjustment methods using SUDAAN (14), a Taylor-Series linearization procedure, which takes into account the complex survey design. The variance estimates were very similar and no significant differences were noted.

Summary

In this paper, we use predicted probabilities from response propensity models to form nonresponse adjustment cells and to minimize the potential for bias in the NHANES III survey estimates. Logistic regression was used to predict response propensity. The sample was then stratified by the response propensity to form approximately equal cell sizes of about 750 sample persons. There were no differences in the weighted estimates from the three response propensity models examined. In addition, a comparison of estimates computed from two alternative strategies for constructing adjustment cells, i.e., stratifying on nonresponse predictor variables or stratifying on response propensity showed no significant differences. Even though there were no differences in the estimates from the two methods, it would probably be prudent to trim the extreme weights from the response propensity method as was done for the weighting class method.

Response propensity modeling for weight adjustment due to unit nonresponse seems appropriate for NHANES III since there is a wealth of interview data available for examination nonrespondents. Thus, a wide range of sociodemographic and health related variables can be used in the adjustment for nonresponse. The response propensity approach is easy to implement, and it allows somewhat greater flexibility than the weighting class procedure since there is no restriction on the number of cells or the type of variables used (numerical versus categorical). In addition, some computation time is saved since no collapsing of cells is required as for the weighting class procedure. Finally, response propensity adjusted weights, like all weights, should be evaluated for excessive variation and trimming of weights done where appropriate (15).

References

1. Ezzati, T., Massey, J., Waksberg, J., Chu, A., and Maurer, K. Sample Design: Third National Health and Nutrition Examination Survey. National Center for Health Statistics. Vital Health Statistics. Series 2, No. 113, 1992.

2. Plan and Operation of the Third National Health and Nutrition Examination Survey, 1988-94. National Center for Health Statistics. Vital Health Statistics. Series 1, No. 32, 1994.
3. Madow, WG., Olkin, I., and Rubin, DB, eds. Incomplete Data in Sample Surveys, Volume 2, New York: Academic Press, 1983.
4. Cox, B. Weighting Survey Data for Analysis. Presentation for the ASA Continuing Education Program. 1991 Joint Statistical Meetings, August, 1991.
5. Kalton, G. and Kaspryzk, D. The Treatment of Missing Survey Data. Survey Methodology, Vol. 12, No. 1, 1-17, Statistics Canada, 1986.
6. Little, RJA. Survey Nonresponse Adjustments for Estimates of Means. International Statistical Review 1986: 54:139-157.
7. Ezzati, T. and Khare, M. Consideration of Health Variables to Adjust Sampling Weights for Nonresponse in a National Health Survey. 1991 Proceedings of the Social Statistics Section of the American Statistical Association, pp. 203-208, 1992.
8. Ezzati, T. and Khare, M. Nonresponse Adjustments in a National Health Survey. 1992 Proceedings of the Survey Research Methods Section of the American Statistical Association, pp. 339-344, 1993.
9. Iannacchione, VG, Milne, JG, and Folsom, RE. Response Probability Weight Adjustments Using Logistic Regression. 1991 Proceedings of the Survey Research Methods Section of the American Statistical Association, pp. 637-641, 1992.
10. Kalton, G. Compensation for Missing Survey Data. Institute for Social Research, University of Michigan, Anne Arbor, 1983.
11. Folsom, RE. Exponential and Logistic Weight Adjustments for Sampling and Nonresponse Error Reduction. 1991 Proceedings of the Social Statistics Section of the American Statistical Association, pp. 197-202, 1992.
12. Judkins, D. and Lo, A. Components of Variance and Nonresponse Adjustment for the Medicare Current Beneficiary Survey. 1993 Presentation of the Survey Research Methods Section of the American Statistical Association, San Francisco, CA.
13. Czajka, JL, Hirabayashi, SM, Little, RJA, and Rubin DB. Projecting from Advance Data Using Propensity Modeling: An Application to Income and Tax Statistics. Journal of Business and Economic Statistics, 10-2, pp. 117-131, 1992.
14. SUDAAN. Research Triangle Institute, Research Triangle Park, 1991.
15. Potter, FJ. A Study of Procedures to Identify and Trim Extreme Sampling Weights. 1990 Proceedings of the Survey Research Methods Section of the American Statistical Association, pp. 225-230, 1991.

Table 1. Comparison of Weighted Estimates Based on Three Response Propensity Models: NHANES III, Phase 1, 1988-91

Measurements	Model 1	Model 2*	Model 3
Height	168.60	168.60	168.59
Weight	74.46	74.49	74.43
Hemoglobin	14.05	14.05	14.05
Serum iron	90.38	90.39	90.34
Total serum cholesterol	206.28	206.25	206.20
Total bone mineral density	0.94	0.94	0.94
Macular degeneration score	0.09	0.09	0.09

*Same as weighting class model.

Table 2. Comparison of Estimates from Two Alternative Nonresponse Adjustment Methods: NHANES III, Phase 1, 1988-91

Measurements	Propensity model	Weighting class model	Percent difference
Height	168.60	168.64	0.025
Weight	74.49	74.51	0.028
Hemoglobin	14.05	14.05	-0.003
Serum iron	90.39	90.44	0.050
Total serum cholesterol	206.25	206.25	0.005
Total bone mineral density	0.94	0.94	0.037
Macular degeneration score	0.09	0.09	-0.185

Table 3. Distribution of Weights from Two Alternative Nonresponse Adjustment Methods: NHANES III, Phase 1, 1988-91

Race/Ethnicity	Minimum	Maximum	Mean	CV
White/other				
Response propensity	4,381	178,444	38,622	53.1
Weighting class	3,758	95,044	38,625	50.5
Black				
Response propensity	3,674	62,617	8,975	42.2
Weighting class	3,606	24,777	9,004	37.0
Mexican-American				
Response propensity	549	15,664	3,432	47.7
Weighting class	609	7,421	3,434	39.2

Table 4. Distribution of Overall Adjustment Factors* from Two Alternative Weighting Methods: NHANES III, Phase 1, 1988-91

Race/Ethnicity	Minimum	Maximum	Mean	CV
White/other				
Response propensity	1.13	2.09	1.71	17.9
Weighting class	0.56	2.62	1.71	15.8
Black				
Response propensity	1.03	2.24	1.47	17.7
Weighting class	0.34	2.33	1.49	17.5
Mexican-American				
Response propensity	0.92	2.05	1.25	16.8
Weighting class	0.45	2.16	1.28	18.4

*Nonresponse and poststratification.