Adjusting for Nonresponse Bias in a Health Examination Survey

MICHAEL L. ROWLAND
RONALD N. FORTHOFER, PhD

Mr. Rowland is a Mathematical Statistician, Division of Health Examination Statistics, National Center for Health Statistics, Presidential Bldg., Rm. 900, 6525 Belcrest Rd., Hyattsville, MD 20782, telephone 301-436-7072. Dr. Forthofer is a Biostatistician residing in Boulder, CO.

Tearsheet requests to Mr. Rowland.

Synopsis

There is a potential for nonresponse bias in most population studies using health examinations. This is true of the Mexican American portion of the Hispanic Health and Nutrition Examination Survey (HHANES), conducted by the National Center for Health Statistics, in which unit nonresponse to the examination accounted for 24 percent of the sample. Even though the full effect of nonresponse can never be really known, ancillary information from the interview sample can be used in an attempt to adjust for bias in estimates.

Two techniques for nonresponse bias adjustment are presented and illustrated using health status level and hypertension status from published studies based on the HHANES of 1982-84. The first approach uses conditional probabilities and the second approach uses direct standardization. The examples examine whether or not an adjustment for socioeconomic status, sex, and age—variables related to both response status and the conditions under study—changes the prevalence estimates of (a) Mexican Americans who report poor, fair, or good health status and (b) hypertension among Mexican Americans.

SURVEYS are playing a prominent role in epidemiologic and health policy studies today. For example, the Health and Nutrition Examination Surveys (HANES) conducted by the National Center for Health Statistics (NCHS) are widely used in epidemiologic research (*1*). There are a number of concerns in the surveys, including sampling variability, measurement error, multiple interviewers and interviewer effect, and nonresponse bias. Andersen and coworkers discuss these issues in detail in "Total Survey Error: Applications to Improve Health Surveys" (*2*).

In this article, we focus on one component of total survey error, nonresponse bias at the unit level, and two methods of adjusting for it. By unit nonresponse, we mean that the sampled person did not respond at all to components of the survey under consideration.

Data from the Mexican American portion of the Hispanic HANES (HHANES) will be used in examples demonstrating the application of these methods. Three separate data collection efforts were used to obtain data from the selected sample participants in HHANES. The first effort collected limited demographic data about the household members. The second survey obtained additional data, for example, medical history data from those in the selected sample who agreed to this longer household interview. Those who agreed to participate at this stage formed the interview sample. The third data collection effort was performed in a mobile examination center and obtained measurements on a wide variety of health and nutritional variables. Those in the interview sample who agreed to this examination formed the examination sample.

Unit nonresponse to the examination—refers to those sample persons who may have provided data at one of the earlier components but did not participate in the examination component—accounted for 24 percent of the sample. The examples examine whether or not an adjustment for socioeconomic status, sex, and age—variables related to both response status and the conditions under study—changes the prevalence estimates of (*a*) Mexican Americans who report poor, fair, or good health status and (*b*) hypertension among Mexican Americans.

Statistical Methodology

When unit nonresponse occurs, there is the possibility of bias. Nonresponse bias is the difference between the population value of a parameter, say the prevalence of a condition denoted by $P(C)$, and the corresponding value for the population of respondents, $P(C_R)$ where R denotes respondent. The expansion of $P(C)$ in terms of respondents and nonrespondents (NR) makes clear the relation of bias to the difference in the prevalences for respondents and nonrespondents as well as to the magnitude of nonresponse. Bias can be expressed as

$$\begin{aligned} \text{bias} &= P(C_R) - P(C) \qquad (1) \\ &= P(C_R) - [P(C_R) \times P(R) + (P(C_{NR}) \times P(NR)] \\ &= P(C_R) \times [1 - P(R)] - P(C_{NR}) \times P(NR) \\ &= [P(C_R) - P(C_{NR})] \times P(NR) \end{aligned}$$

If there is no difference between the prevalence rates for respondents and nonrespondents, then there is no bias due to nonresponse. If the proportion of nonrespondents is small, then the bias will likely be small as well, and the differences between the estimates of prevalence should also be small.

Nonresponse weighting adjustments. Survey organizations typically attempt to adjust for this potential unit nonresponse bias by forming adjustment cells created by cross-tabulating a few demographic and socioeconomic variables. Demographic and socioeconomic variables are often used because (*a*) even if the target respondent does not participate, it may be possible to obtain information on these items; and (*b*) response status often is highly related to these variables. The survey organization then is able to provide the target number of persons in each of the adjustment cells as well as the actual number of respondents in these cells. The adjustment consists of multiplying the sample weight for each respondent in an adjustment cell by the ratio of the weighted target sample size to the weighted actual sample size in the adjustment cell (*3–7*).

In using this approach, one makes the assumption that the nonrespondents in the adjustment cell would have responded similarly to the respondents in that cell. Regardless of the method of adjustment employed by the survey group, it is necessary at some stage to make this assumption about the similarity of respondents and nonrespondents.

If there is substantial nonresponse, it is important to examine the data for evidence of bias. However, it is difficult to do this unless there is information available about the target population. Examples of such investigations are a previous study of nonresponse in the second National Health and Nutrition Examination Survey (*8*) and a similar study by the authors of nonresponse in the HHANES (in preparation). Comparison groups used in the present study were the HHANES Mexican American interview sample and the subpopulation of Mexican Americans in the National Health Interview Survey (NHIS).

'Suppose that in a survey designed to estimate the prevalence of disease, there is a differential response rate by age. Because of the possible nonresponse bias associated with age, it would be inappropriate to report a crude rate based on this sample. However, if one assumes that the age-specific prevalence rates are unbiased, they are of interest and could be reported.'

Some socioeconomic variables were associated with response status at the examination stage and at the interview stage, although the marginal and two-way distributions of these variables in the HHANES examination and interview samples were very similar. Comparison of these same variables from the HHANES examination and interview samples to data from the NHIS suggested the possibility of a bias in the distribution of the perceived health status, possibly due to socioeconomic status.

Adjusting for nonresponse bias

Probability approach. This technique, used previously at NCHS (*9*), incorporates a variable(s) to modify the sample estimate of the population parameter, for example, the prevalence of some condition C. This variable(s) should be related to both the response status and the prevalence of condition C.

Suppose that the prevalence of hypertension is to be estimated. The prevalence is known for the respondents (R) but not for the nonrespondents (NR). The proportions of respondents, $P(R)$, and of nonrespondents, $P(NR)$, are also known. For example, using the HHANES data, the prevalence of hypertension for those examined will be known, but it is unknown for those who were interviewed but not examined.

Using the examination data, it is found that the prevalence of hypertension is related to the levels of a variable V which was obtained during the household interview. If the levels of V are also related to the examination response status, then V can be used to adjust the prevalence estimate for C in the following fashion:

$$P(C) = \left[\sum_{i=1}^{I} P(C \mid V_{iR}) \times P(V_{iR}) \right] \times P(R) \qquad (2)$$

$$+ \left[\sum_{i=1}^{I} P(C \mid V_{iR}) \times P(V_{iNR}) \right] \times P(NR)$$

The first term in brackets is $P(C_R)$. The second bracketed term is $P(C_{NR})$ under the assumption that the relation between the prevalence of condition C and variable V is the same for respondents and nonrespondents. $P(V_{iR})$ denotes the proportion of the respondents at level i of V and $P(V_{iNR})$ denotes the proportion of the nonrespondents at level i of V. Note that $P(C \mid V_{iR})$, the conditional probability of the condition C given level i of variable V, is used for both the respondents and nonrespondents, as was mentioned earlier. The examination data provide an estimate of this relation. Given this assumption, all the pieces of this equation can be estimated from the data, and it is possible to obtain an adjusted estimate of the prevalence of C.

This probability approach is equivalent to the adjustment cell approach. Equation 2 simply shows another formulation for the adjustment cell method, a formulation that may be more familiar to people not versed in sample survey methodology. It makes clear the assumption of the equality, for respondents and nonrespondents, of the relation between the prevalence of C and the variable V.

This procedure is not limited to using a single variable because the levels of V could be combinations formed by cross-tabulating two or more variables. The selection of V is an important issue. Ideally, the researcher knows what variables would be appropriate candidates for V. If there is little guidance for the selection of V, then statistical methods can be employed. For example, a logistic regression could be run with the condition C as the dependent variable and a number of the interview questions could be selected to be the independent variables. If any of the independent variables are identified as being related to C, then they could be used as independent variables in a second logistic regression with the response status to the examination stage as the dependent variable. The variable(s) related to both C and the response status to the examination are candidates for V.

Direct standardization approach. Suppose that in a survey designed to estimate the prevalence of disease, there is a differential response rate by age. Because of the possible nonresponse bias associated with age, it would be inappropriate to report a crude rate based on this sample. However, if one assumes that the age-specific prevalence rates are unbiased, they are of interest and could be reported. If the correct marginal age distribution is known, then direct standardization (*10*) could be used to obtain a prevalence rate adjusted for the nonresponse. Direct standardization is a term from demography and epidemiology, and it is equivalent in this context to the term "post-stratification" used by sampling statisticians.

Sensitivity to assumptions. As previously mentioned, it is necessary in adjusting for nonresponse bias to make assumptions about the similarity of respondents and nonrespondents. For both adjustment approaches used in this paper, the relationship between the condition C and the adjustment variable V is assumed to be the same in respondents and nonrespondents. Although this assumption cannot be verified, if it were incorrect, an error in either direction could have been introduced in the final bias-adjusted estimate. An example of a sensitivity analysis of bias estimation to the assumptions of the analysis is given by Hadden and Harris (*9*).

Empirical Examples

Survey design. In the HHANES conducted from July 1982 through December 1984, the following three groups of Hispanics were examined: Mexican Americans residing in five southwestern States (Arizona, California, Colorado, New Mexico, and Texas); Cuban Americans residing in Dade County, FL; and Puerto Ricans residing in the New York City metropolitan area, including parts of New York, New Jersey, and Connecticut. Data were collected using a design that called for a stratified, multistage, probability sample of persons ages 6 months through 74 years (*3*). Only the Mexican American sample of persons ages 20–74 of Mexican origin or ancestry is used in the examples presented in this report.

Variables and definitions. As was mentioned previously, the HHANES included household interviews and a physical examination. The variables of interest from the household interviews were the self-reported health status of the sample person and the total family income. These variables were of interest due to the differences in their distributions between the HHANES and the NHIS during the same period. Other variables used in the examples and their definitions come from the physical examination, and they are as follows:

Blood pressure. Systolic (first phase) and diastolic (fifth phase) blood pressure were measured to the nearest even digit using a standard mercury sphygmomanometer. Two blood pressure measurements were taken on one occasion with the patient seated, and the average of the two readings was used for the analysis presented in this paper.

Hypertension. The definitions used by Pappas and coworkers (*11*) also were used in this paper. Persons were considered hypertensive if they met one of the following conditions: (*a*) diastolic blood pressure of 90 millimeters or more of mercury (mm Hg); (*b*) systolic blood pressure of 140 or more mm Hg; (*c*) currently using anti-hypertensive medication regardless of blood pressure value.

Statistical analysis. Weighted data are used in both of the adjustment approaches. The estimation of crude prevalence rates among respondents is done using the HHANES Mexican American examination sample weights. These weights are the reciprocal of the probability of selection with adjustments for nonresponse and noncoverage and a final post-stratification to U.S. Census Bureau population estimates. The computation of the component factors of the adjusted estimates was done using basic weights, that is, the reciprocal of the probability of selection. Note that odds ratios shown in table 1 were computed using the examination weights, whereas conditional prevalence estimates shown in table 2 were computed using basic weights. Thus, due to the use of different weights, the odds ratios shown differ slightly from the odds ratios that would result using basic weighted prevalence estimates.

The complex survey design used in the HHANES tended to increase the estimated variance of prevalence estimates over that which would have been obtained through simple random sampling (*12*). An increased variance due to the survey design (that is, design effect) would have led to wider confidence intervals. Since the conclusion of this paper (that there was no significant difference between unadjusted and bias-adjusted prevalence estimates) would have been unchanged if the larger confidence intervals (CIs) had been used, the simple random sample CIs are presented.

All data analyses were done using SAS procedures (*13*) or programs accessible through SAS (*14*).

Table 1. Relation of selected health status and disease prevalence conditions to the predictor variable total family income, according to sex and age for Mexican Americans 20–74 years, Hispanic Health and Nutrition Examination Survey, 1982–83

Selected variables, sex, and age (years)	Examination sample size	Total family income	
		Odds ratio	95 percent CI[1]
Health status:			
Women 20–44	1,121	2.03	1.56–2.64
Men 20–44	886	2.76	2.08–3.66
Women 45–74	670	2.63	1.68–4.11
Men 45–74	524	2.10	1.38–3.18
Elevated blood pressure:			
Women 20–44	1,095	1.25	0.57–2.74
Men 20–44	861	0.64	0.40–1.02
Women 45–74	655	1.70	1.12–2.59
Men 45–74	517	1.12	0.77–1.64
Hypertension:			
Women 20–44	1,095	1.39	0.72–2.66
Men 20–44	861	0.66	0.41–1.04
Women 45–74	655	1.42	0.98–2.06
Men 45–74	517	1.25	0.86–1.80

[1]CI = confidence interval.

NOTE: Odds ratio is the ratio of the odds of good, fair, or poor self-reported health status (or elevated blood pressure or hypertension) for those with total family income of less than $20,000 to the odds for those with total family income of at least $20,000.

Results

The object in the first example was to estimate the "true" prevalence of self-reported poor, fair, or good health status, rather than excellent or very good health status, for 20–74-year-old Mexican American persons in the HHANES sample, given that nearly 30 percent of all persons in the sample were not examined. The two methods of approximating this prevalence involved utilizing the relationship between self-perceived health status and other variables in the data set.

Sex, age, and total family income were selected as the variables whose cross-tabulation form the levels of V for this example. These variables were associated with nonresponse to the examination (unpublished findings), and tables 1 and 2 show that they are also related to self-reported health status. The odds ratios for the relation of the self-reported health status and income, with $20,000 as the cutpoint, range from 2.03 to 2.76 for the different age and sex combinations. For

Table 2. Bias estimation for prevalence of self-perceived health of "poor," "fair," or "good" for Mexican Americans 20–74 years in the Health Examination Survey sample: HHANES,[1] 1982–83

Sex, age, and income	Examination sample size	$P(C \mid Vir)$	$P(Vir)$	$P(C \mid Vir)$	$P(Vinr)$
Women					
20–44 years with income less than $20,000....	741	0.7180	0.226	0.7180	0.152
20–44 years with income $20,000 or more.....	380	0.5690	0.124	0.5690	0.097
45–74 years with income less than $20,000....	499	0.8870	0.108	0.8870	0.086
45–74 years with income $20,000 or more.....	171	0.7427	0.038	0.7427	0.036
Men					
20–44 years with income less than $20,000....	547	0.7059	0.229	0.7059	0.318
20–44 years with income $20,000 or more.....	339	0.4734	0.148	0.4734	0.180
45–74 years with income less than $20,000....	336	0.8260	0.081	0.8260	0.083
45–74 years with income $20,000 or more.....	188	0.6978	0.046	0.6978	0.048
Adjusted prevalence: sex-, age-, income-adjusted estimate = 68.50					

[1]Hispanic Health and Nutrition Examination Survey.

NOTE: Adjusted prevalence $= 100 \times [[P(C \mid Vir) \times P(Vir)] \times P(R) + [P(C \mid Vir) \times P(Vinr)] \times P(NR)]$

where,

C = disease condition.

V = adjustment variable(s).

R = respondents.

NR = nonrespondents.

$P(R)$ = proportion responding = 0.7024.

$P(NR)$ = proportion not responding = 0.2976.

$P(C \mid Vinr)$ = proportion of nonrespondents at level Vi with disease condition C.

$P(C \mid Vir)$ = proportion of respondents at level Vi with disease condition C.

$P(Vir)$ = proportion of respondents at level i of V.

$P(Vinr)$ = proportion of nonrespondents at level i of V.

example, for women ages 20–44 years with a total family income less than $20,000, the odds of having a self-reported health status of good, fair, or poor was about twice (2.03) that of 20–44-year-old women whose total family income was at least $20,000.

Table 2 shows clearly the effect of sex, age, and income with those who are male, younger, and making $20,000 or more having the smaller percentages of those with poor, fair, or good self-reported health status and, thus, the larger percentages of excellent or very good self-reported health status. In the NHIS, the percentage with poor, fair, or good self-reported health status was approximately 50 percent, substantially less than the approximately 70 percent shown in table 3 for HHANES participants.

The two methods used to derive adjusted estimates of the prevalence in the total population were the (*a*) conditional probability approach and the (*b*) direct standardization approach. For (*a*), data internal to the HHANES were used to make the adjustments. The adjusted prevalence was computed using equation 2 and is illustrated in table 2. The overall estimate of the proportion with poor, fair, or good self-perceived health status from the examination was 68.77 percent, and is shown in table 3 along with the nonresponse adjusted estimate of 68.50 percent.

There is little difference in these estimates because the age, sex, and income groups with the greatest differences in representation between respondents and nonrespondents (young women and men with family incomes less than $20,000) have approximately the same proportion of poor, fair, and good self-reported health status (71.80 and 70.59 percent).

In the direct standardization approach, information external to the HHANES was used to make bias adjustments. A single summary bias-adjusted prevalence estimate was obtained by applying the schedule of sex-age-income specific prevalence rates of self-perceived health status of poor, fair, or good estimated from the HHANES examination sample against the sex-age-income population proportional distribution of the subpopulation of Mexican Americans in the NHIS for the years 1982 through 1984 combined. The rationale for using the NHIS distribution was that it was arguably more representative of the target population than the HHANES due to its assumed smaller nonresponse rate. This adjusted prevalence estimate of 67.7 percent is shown in table 3. All of these adjusted estimates are contained within the 95 percent CI for the unadjusted estimate.

The second example examines the prevalence of hypertension in Hispanics, an issue that has received attention recently (*11,15–17*). Pappas and coworkers (*11*) and Geronimus and coworkers (*15*) used data from HHANES to estimate the prevalence of hypertension for Hispanics. Both studies point out that the rates are less than those for anglos and blacks. Pappas and coworkers suggest that this finding is reasonable, whereas Geronimus and coworkers argue, based on comparisons with Hispanic women in NHANES II and other studies,

Table 3. Crude and age-, sex-, and income-adjusted estimates for selected health status and disease prevalence conditions for Mexican Americans 20–74 years: HHANES, 1982–83

Selected variables	Crude estimates (percent)	95 percent confidence interval	Bias-adjusted estimates	
			Probability approach (percent)	Standardization approach (percent)
Health status of "poor," "fair," or "good"	68.77	67.16–70.38	68.50	67.70
Elevated blood pressure	13.40	12.21–14.59	13.38	12.85
Hypertension	16.36	15.06–17.66	16.23	15.72

that HHANES may not provide reliable prevalence estimates for Hispanic women. Johnson and Woteki (*16*) point out that there are differences in the two studies which make exact comparisons problematic and that it is also difficult to compare these results with data from other studies because of differing conditions between studies. Caralis (*17*) reviewed the data on hypertension in Hispanics and found wide variability in the prevalence rates, although with several studies suggesting that the prevalence is higher in Hispanics than in anglos.

The purpose of this second example is to examine whether or not a nonresponse adjusted prevalence estimate is much different from the unadjusted estimate computed using the same definition of hypertension as used by Pappas and coworkers and Geronimus and coworkers. The unadjusted examination prevalence estimate of hypertension is 16.36 percent compared with the conditional probability adjustment value of 16.23 percent and the direct standardization value of 15.72 percent. Although there may be some underrepresentation of the higher income Hispanics in HHANES, it does not appear to play a major role in the estimate of the prevalence of hypertension. The direct standardization method produces the larger change, a relative reduction of almost 4 percent (an absolute reduction of 0.64 percentage points) in the prevalence estimate, which is in the opposite direction from that suggested by Geronimus and coworkers.

Discussion

It is common practice to modify the basic sampling weights to adjust for potential unit nonresponse bias in surveys. This is usually accomplished through the use of adjustment cells created via the cross-tabulation of a few demographic and socioeconomic variables. However, the researcher sometimes finds differential nonresponse with regard to other variables that also have an important relation to the analytic variable of interest. In other cases there may be evidence, perhaps from a comparison

'. . . there may be evidence, perhaps from a comparison with another survey, that nonresponse weighting has not adequately adjusted for noncoverage or nonresponse in certain segments of the population. Examples of this include the total family income and education of head of household variables, which differed between the Mexican American component of the HHANES and the Mexican Americans in the NHIS . . . .'

with another survey, that nonresponse weighting has not adequately adjusted for noncoverage or nonresponse in certain segments of the population. Examples of this include the total family income and education of head of household variables, which differed between the Mexican American component of the HHANES and the Mexican Americans in the NHIS (unpublished findings).

In lieu of reporting only stratum specific results in every analysis using the potentially biasing nonresponse variables or readjusting the survey weights, this study provides procedures for adjusting the estimates. For the two situations mentioned, the methods developed and illustrated in this paper dealt with the former problem using a probability approach, and with the latter problem using direct standardization. When there is good external data, as in this example where there is thought to be little nonresponse for the Mexican Americans in the NHIS, the direct standardization approach may be the adjustment method of choice. However, it is unusual that good external data will be available and, hence, the probability approach can still be used to adjust for unit nonresponse.

Little (*18*) examines the bias and mean square error of three methods—weighting by the inverse of the response rate within the adjustment cell, which is equivalent to our probability method; post-stratification on known cell counts, or what we have labeled as direct standardization; and mean imputation within the adjustment cells, adjusting means for unit nonresponse.

Little's simulation results are of interest. The mean square error results have to be interpreted with caution since they are based on simple random sampling. Basically, Little's results supports the use of direct standardization (post-stratification) when it can be used. The results also suggest that the use of the variable *V* for creation of new adjustment cells is promising and can help to control mean square error. Whether direct standardization or the probability method is used, the simulation results suggest that they will reduce mean square error compared with using the unadjusted estimators.

There are also methods for adjusting for item nonresponse, that is, a person participates in the survey sample but fails to supply data for a few of the requested items. Rubin (*19*) provides one approach, a Bayesian approach to adjusting for item nonresponse.

The two examples presented reinforce the use of data from HHANES. The examples selected were thought to be some of the worse situations, that is, cases where there was potential for substantial bias because of the large difference in socioeconomic status between HHANES and NHIS. However, although these socioeconomic differences were present, they seemed to have little to no effect on the estimated prevalence of poor, fair, or good self-perceived health status or on the prevalence of hypertension. These examples support but do not provide for the carte-blanche use of HHANES. As with any data set, care must be exercised in assessing the potential biasing effects of nonresponse for each analysis done.

References

1. Selective bibliography of HHANES publications. Am J Public Health 80: 71-72, December 1990.
2. Andersen, R., Kasper, J., Frankel, M. R., and associates: Total survey error: applications to improve health surveys. Jossey-Bass, Inc., San Francisco, CA, 1979.
3. National Center for Health Statistics: Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982-84. Vital Health Stat [1], No. 19. DHHS Publication No. (PHS) 85-1321. U.S. Government Printing Office, Washington, DC, September 1985.
4. Kalton, G.: Compensation for missing survey data. Institute for Social Research, University of Michigan, Ann Arbor, 1983.
5. Gonzalez, J. F., Jr., Ezzati, T. M., Lago, J., and Waksberg, J.: Estimation in the Southwest component of the Hispanic Health and Nutrition Examination Survey. *In* Proceedings of the Survey Research Methods Section of the American Statistical Association, August 1985, pp. 233-237.
6. Landis, J. R., Lepkowski, J. M., Eklund, S. A., and Stehouwer, S. A.: A statistical methodology for analyzing data from a complex survey, the First National Health and Nutrition Examination Survey. Vital Health Stat [2], No. 92, DHHS Publication (PHS) 82-1366. National Center for Health Statistics. U.S. Government Printing Office, Washington, DC, September 1982.
7. Levy, P. S., and Lemeshow, S.: Sampling for health professionals. Wadsworth, Inc., Belmont, CA, 1980.
8. Forthofer, R. N.: Investigation of nonresponse bias in NHANES II. Am J Epidemiol 117: 507-515 (1983).
9. Hadden, W. C., and Harris, M. I.: Prevalence of diagnosed diabetes, undiagnosed diabetes, and impaired glucose tolerance in adults 20-74 years of age, United States, 1976-80. Vital Health Stat [11], No. 237, DHHS Publication No. (PHS) 87-1687. National Center for Health Statistics. U.S. Government Printing Office, Washington, DC, February 1987.
10. Fleiss, J. L.: Statistical methods for rates and proportions. John Wiley & Sons, Inc., New York, NY, 1981.
11. Pappas, G., Gergen, P. J., and Carroll, M.: Hypertension prevalence and the status of awareness, treatment, and control in the Hispanic Health and Nutrition Examination Survey (HHANES), 1982-84. Am J Public Health 80: 1431-1436 (1990).
12. Kovar, M. G., and Johnson, C.: Design effects from the Mexican American portion of the Hispanic Health and Nutrition Examination Survey: a strategy for analysts. *In* Proceedings of the Section on Survey Research Methodology of the American Statistical Association. American Statistical Association, Washington, DC, 1986, pp. 396-399.
13. SAS Institute Inc.: SAS user's guide, Ed. 5. SAS Institute, Inc., Cary, NC, 1982.
14. Research Triangle Institute: SUDAAN: software for survey data analysis, version 5.30. Research Triangle Park, NC, 1991.
15. Geronimus, A. T., Neidert, L. J., and Bound, J.: A note on the measurement of hypertension in HHANES. Am J Public Health 80: 1437-1442 (1990).
16. Johnson, C. L., and Woteki, C. E.: The art and science of interpreting survey data. Am J Public Health 80: 1427-1429 (1990).
17. Caralis, P. V.: Hypertension in the Hispanic-American population. Am J Med 88 (supp 3B): 9S-16S (1990).
18. Little, R. J. A.: Survey nonresponse adjustments for estimates of means. Int Stat Rev 54: 139-157 (1986).
19. Rubin, D. B.: Formalizing subjective notions about the effect of nonrespondents in sample surveys. J Am Stat Assoc 72: 538-543 (1977).