An Evaluation of Weighting Methodology in an RDD Survey with Multiple Population Controls

Kennon R. Copeland[1], Meena Khare[2], N. Ganesh[1], Zhen Zhao[3], Abera Wouhib[2]
[1]NORC at the University of Chicago, 4350 East-West Hwy, Ste 800, Bethesda, MD 20814
[2]National Center for Health Statistics, CDC, 3311 Toledo Road, Hyattsville, MD 20782
[3]National Center for Immunizations and Respiratory Diseases, CDC, 1600 Clifton Road, Atlanta, GA 30333

Abstract
To control potential noncoverage bias, ratio adjustment methods are often used to adjust survey population estimates to represent relevant subgroups in the target population. The National Immunization Survey (NIS)—a nationwide, list-assisted RDD survey fielded by the NORC for the Centers for Disease Control and Prevention—monitors the vaccination rates of children between the ages of 19 and 35 months and adolescents between the ages of 13 and 17 years. The NIS uses various ratio adjustment methods based on multiple population controls and distributions to adjust survey estimates. The NIS also utilizes the Keeter adjustment method using information on telephone service interruptions to adjust for noncoverage of nonlandline telephone households. This research evaluates the potential impact of adjusting across multiple population controls on survey weights, estimates, and variances, and seeks to identify a best strategy for refinements.

Key words: ratio adjustment, noncoverage bias, variance

1. Introduction[1]

Random-digit-dial (RDD) telephone surveys are a cost-effective and quick method for conducting household surveys. However, a major weakness of an RDD survey is that it only covers the households with landline telephone, leaving non-landline (phoneless and cell-only) households out of coverage. Based on the National Health Interview Survey (NHIS), the proportion of non-landline households has increased from 2.8% in 2003 to 22.1% in 2008, with 21.1% of children under age 18 living in non-landline households in 2008 (Blumberg and Luke, 2009). The increase over time is due almost exclusively to the increase in cell-only households, with 20.2% of households being cell-only in 2008 (and 0.9% with no phone), and 18.7% of children under age 18 living in cell-only households in 2008 (and 2.4% in no phone households). The proportion of young children living in non-landline households is even greater. Table 3 of this paper shows that, based on data from the 2007 NHIS, an estimated 26.5% of children 19-35 months of age (the eligible age range for the NIS) were living in non-landline households; of those 22.1% were in cell-only households and 4.4% were in phoneless households. With such an increasing

[1] "The findings and conclusions in this paper are those of the author(s) and do not necessarily represent the views of Centers for Disease Control and Prevention."

rate of noncoverage, the potential for coverage bias in RDD surveys is being questioned more extensively.

While frame issues certainly contribute to the majority of coverage issues in RDD surveys, there are other potential causes for noncoverage: differential participation and inability to identify all eligible persons in a sample household.

The common approach to addressing noncoverage in survey data processing is to apply a ratio adjustment during sample weighting to control weighted sample counts to population totals for characteristics known or believed to be correlated with the variables of interest. In this paper, we investigate the ratio adjustments applied in the case of the National Immunization Survey (NIS), which is an RDD survey that monitors the immunization coverage among young children and adolescents. We try to assess the impact of the ratio adjustments on estimates and corresponding variances. We also compare the characteristics of the households by telephone status (landline without interruption, landline with interruption, cell-only, and phoneless) to explore the possibility of developing an improved adjustment procedure by making separate adjustments for non-landline (cell-only and phoneless) households. For this latter comparison, since the NIS does not cover non-landline households, we use data from the NHIS.

1.1 The National Immunization Survey

The NIS has been conducted quarterly since 1994 by the Centers for Disease Control and Prevention (CDC), to estimate the vaccination coverage rates among children aged 19 to 35 months in the U.S. within geographic areas (called estimation areas) consisting of 50 states, the District of Columbia, and several sub-state areas. The NIS collects vaccination data on the following childhood vaccines: diphtheria, tetanus toxoids and pertusis (DTaP) vaccine; poliovirus (polio) vaccine; measles, mumps and rubella (MMR) vaccine; *Haemophilus influenza* type b (Hib) vaccine; hepatitis B (Hep B) vaccine; varicella zoster (chicken pox) vaccine; pneumococcal vaccine; hepatitis A (Hep A) vaccine; influenza vaccine; and rotavirus vaccine. The NIS uses a two-phase survey design where the first phase is a RDD survey that identifies the households with age-eligible children and collects information on vaccinations and vaccination providers of the eligible children. In the second phase a mail survey of providers, called the provider record check (PRC), collects detailed vaccination histories for the children for whom the RDD-phase interview was complete and consent to contact providers was received. In 2007, the NIS included 24,807 children with complete household interviews and 17,017 children with adequate provider data.

The NIS weighting methodology starts with a base weight reflecting the probability of selection. Nonresponse adjustments for telephone number resolution, household screening, and household interview are performed on the quarterly samples. The quarterly samples are then combined to create an annual sample, with the quarterly weights adjusted proportional to the completed interviews obtained from each quarter. Annual weight adjustment of the NIS sample includes the following ratio adjustments to account for coverage error (see Smith, *et al.*, 2005), in addition to another nonresponse adjustment for provider data collection that occurs after the third listed ratio adjustment:

1) Keeter adjustment for undercoverage of children in non-landline telephone households;
2) Simple post-stratification of weights for children with completed household interviews to population controls at the mother's race/ethnicity x mother's education x child's age group level;
3) Raking of weights for children with completed household interviews to mother's race/ethnicity and education, and child's age and gender; and
4) Raking of weights for children with adequate provider data to mother's education, child's race/ethnicity, age, gender, and first-born status, and provider response propensity quintile.

2. Analysis of Current Weights

The NIS weighting methodology applies truncation at each ratio adjustment stage in an attempt to control for large deviations among survey weights. In addition, truncation is carried out in the Keeter adjustment to control for large deviations between the adjustments applied to the telephone service interruption/no interruption sample. However, the weighting consists of many stages, each making adjustments to the survey weights. There is currently no control over the cumulative adjustment which can be made to the weight for an individual sample unit.

2.1 Ratio Adjustment Issues

Ratio adjustment is carried out to control for coverage bias. In the case of the NIS, coverage bias can result from: 1) non-coverage of the cell-only and non-telephone populations (and for which the Keeter adjustment is applied); 2) sample imbalances relative to population characteristics of interest; and 3) under-ascertainment (or the failure to identify eligible children in households for which the screening interview was completed).

However, while ratio adjustment can control the potential for coverage bias, the resultant weight changes can serve to increase estimate variance. The current NIS ratio adjustments vary in terms of their impact on survey weights. As shown in Table 1, nonresponse adjustments have relatively small impact, due to the relatively high response rates achieved in the survey. By contrast, the Keeter adjustment results in a median impact of doubling the survey weights from the prior weighting step, with the 5^{th} to 95^{th} percentile range being 1.6 to 3.6 (bolded entries in Table 1, immediately below "Annual Level Household Weights"). The remaining ratio adjustments have relatively small impact on the weights.

The net result of the NIS weighting steps is to increase the annual household weights by a median factor of 2.4, with the 5^{th} to 95^{th} percentile range being 1.3 to 7.8 (bolded entries in lower right hand corner of Table 1). These results point out the potential for variance impact issues associated with the ratio adjustment steps.

Adjusting sample estimates to agree with population controls will only reduce coverage bias in survey estimates if the population controls are correlated with survey variables of interest. The factors utilized in the NIS ratio adjustments include child's race/ethnicity, age group level, gender, and first-born status, and mother's race/ethnicity and education level. As part of this research, the relationship between these and other population controls was examined for potential consideration in refining the ratio adjustments.

2.2 Weighted Distributions

Weighted sample distributions for selected population controls were examined by comparing the 2007 NIS and the 2007 NHIS (as 2007 was the most recently available complete year available for both surveys at the time the analysis was conducted) for children 19-35 months of age (Tables 2-4). For the NIS, results were examined by telephone service interruption status (service interruption of one week or longer in the prior 12 months, no service interruption), which is used in the Keeter adjustment. For the NHIS, results were examined by telephone status (landline (with no cell phone, with some cell phone, with mostly cell phone), no landline (cell phone, no phone)) and, for sample in a landline household, telephone service interruption status. NHIS survey estimates are post-stratified to sex, age, and race/ethnicity population controls, with the age group encompassing the NIS eligible population being under age 5.

As expected, sample distributions by race/ethnicity are comparable for the NIS and the NHIS (but are not equivalent due to the differences in age groups used in sample weighting), as seen by examining Table 2. Although significance tests were not carried out, the estimated proportions of all NIS children as compared to estimated proportions of NHIS children in landline households were between five and fifteen percentage points:

a) greater for mother's age $\geq$ 30, mother's education of 12 years, and households with five or more persons; but
b) less for mother's age < 30, mother's education > 12 years, and households with three or four persons.

Note that these variables are not used in NHIS final weighting, while only mother's education is used in NIS final weighting. Differences in estimated proportions for race/ethnicity (used in both NIS and NHIS final weighting) were all less than five percentage points. NHIS distributions for income to poverty ratio contained more than 25 percent missing/don't know, so comparisons were not made for this characteristic.

Also of interest are the relationships with telephone characteristics based upon NHIS data, as shown in Table 3. Although significance tests were not carried out, the estimated proportions of children living in non-landline households as compared to those for children in landline households:

1) were between 30 and 35 percentage points:
   greater for mother's age < 30 years; but
   less for mother's age $\geq$ 30 years;
2) were between 15 and 25 percentage points:
   greater for mother's education < 12 years; but
   less for mother's education > 12 years;
3) were between 10 and 20 percentage points:
   greater for household income to poverty ratio < 1.0; but
   less for household income to poverty ratio $\geq$ 4.0;
4) were between five and fifteen percentage points:
   greater for child' race/ethnicity of Hispanic: but
   less for child's race/ethnicity of non-Hispanic White and non-Hispanic Black; and
5) were twice as large for households with two persons.

Finally, examining distributions by telephone service interruptions, as provided in Table 4, it can be seen that the estimated distributions of NIS children in households with a telephone service interruption often differ from those of NHIS children in cell-only and phoneless households by ten percentage points or more. The exceptions are for race/ethnicity, where the distribution for NIS telephone service interruption is within ten percentage points of the distribution for NHIS cell-only, and for mother's age, where the distribution for NIS telephone service interruption is within ten percentage points of the distribution for NHIS phoneless. By these measures, the NIS no telephone service interruption distributions provide roughly the same level of comparability to the NHIS cell-only distributions as do the NIS telephone service interruption distributions, although the same is not true for the NHIS phoneless distributions. Thus, it appears that telephone status and telephone service interruption alone could not be used to adjust for non-coverage of non-landline households.

While the NIS landline telephone service interruption sample is poorer and mothers have lower education and age, it is also true that they do not mirror the population on these characteristics and on race/ethnicity as measured by the NHIS, as shown in Table 4. These findings suggest the Keeter adjustment, which merely adjusts sample weights for telephone service interruption sample to account for non-landline telephone households, may not be sufficiently focused to address potential differences in the sample and population. This is an important consideration given that telephone service interruptions account for roughly 6% of the 2007 NIS sample and, based on the 2007 NHIS, children in non-landline households make up approximately 26% of the total population of 19-35 month olds.

3. Importance of Telephone Characteristics in Predicting Health Status

It is apparent that NIS must make adjustments to the survey weights to account for undercoverage of non-landline telephone households. Currently, adjustment is made through the Keeter approach, using information on telephone service interruption status.

An important question is whether telephone characteristics (such as landline vs cell or telephone service interruption) are important correlates of the NIS variables of interest, or whether geographic and sociodemographic characteristics are the key correlates. While NIS data can be used to look at the importance of telephone interruption status, it cannot be used to answer the broader question which relates to cell-only and no phone.

3.1 Modeling of Health Status Variables

NHIS and NIS weighted data were examined in an attempt to determine the importance of telephone status for the NIS. This was done using two related models, the primary difference being the telephone characteristics used.

NHIS: $\phi_1\left(E\{Y_1\}\right) = \alpha_1 + X_1\beta_1 + Z\theta_1$

NIS: $\phi_2\left(E\{Y_2\}\right) = \alpha_2 + X_2\beta_2 + Z\theta_2$

where the Y's represent the health variables, X's represent the telephone characteristic variables, and Z's represent geographic and sociodemographic variables. The ϕ's represent functions, the α's represent intercepts, and the β's and θ's represent coefficients. For purposes of this analysis, logistic regression models were selected for use as the Y variables used are bivariate.

The health variables used for NHIS were: flu shot within the prior 12 months; history of asthma, and prior episode of chicken pox. Those used for NIS were: flu shot within the prior 12 months and up to date (UTD) status for 4:3:1:3:3:1[2] vaccination series.

The telephone characteristic used for NHIS was telephone status (landline only, cell-only, cell-mostly, landline mostly, no phone), while for NIS it was telephone service interruption status (service interruption of one week or longer in prior 12 months, no service interruption).

The sociodemographic variables used were region (4 Census regions), MSA status (MSA, non-MSA), child gender, race/ethnicity (Hispanic, Black, non-Hispanic, White non-Hispanic, Other non-Hispanic) and age (one year old, two years old, three years old, 4 years old), mother's education level (less than high school, completed high school, more than high school) and age (<25, 25-29, 30-34, 35+), household income to poverty ratio (<1.0, 1.0-1.99, 2.0-3.99, 4.0+), insurance coverage status (coverage, no coverage), size (2 persons, 3 persons, 4 persons, 5+ persons), and (for NIS) Vaccines for Children program status (participant, non-participant).

3.2 Analysis of Results

A stepwise logistic regression was used to identify the significant variables for predicting the healthcare variable. NHIS models utilized sample persons aged one to four years to provide sufficient sample sizes, while NIS models utilized sample persons aged 19 to 35 months. Normalized sample weights were used to account for the complex sample design and differential probability of selection for both the NHIS and the NIS. Results are provided in Tables 5 and 6.

For both the NHIS and the NIS models, for all health variables, telephone characteristics were not significant given the presence of the sociodemographic variables.

When looking at the common health variable, flu shot within the prior 12 months, both NHIS and NIS identified child's age and household income as significant factors. Although not presented in the tables, lower household incomes were associated with lower vaccination coverage rates, and 24-35 months olds evidenced lower vaccination coverage rates than under 24 month olds for both NHIS and NIS. The NHIS model also included region, MSA status, child's race/ethnicity, and household insurance coverage. The NIS model also included mother's education level.

When looking at a key NIS variable of interest, UTD status for the 4:3:1:3:3:1 vaccination series, the NIS model included both child and mother's age, and household insurance coverage, size, and VFC status.

[2] 4 or more doses of DTaP, 3 or more doses of poliovirus vaccine, 1 or more doses of any MMR, 3 or more doses of Hib, 3 or more doses of HepB, 1 or more doses of varicella vaccine.

3.3 Analysis of NIS Predictor Variables

NIS collects additional sociodemographic variables beyond those used in the NIS/NHIS modeling analysis. The additional variables are child's first-born status, mother's marital status, and number of children in the household.

A full set of vaccination series were used as well: 4+DTaP[3]; 3+Polio[4]; 1+MMR[5]; 3+Hib[6]; 3+HepB[7]; 1+Var[8]; 3+PCV[9]; 4:3:1[10]; 4:3:1:3[11]; 4:3:1:3:3[12]; and 4:3:1:3:3:1 (the vaccine series used in the NIS and NHIS model comparison).

A stepwise logistic regression was again used to identify the significant variables for predicting vaccination coverage. Results are provided in Table 7, with the numbers provided in each column representing the order of entry for the variable into the stepwise logistic regression model for the corresponding vaccination. A blank for a variable indicates the variable was not significant for that vaccination.

As can be seen, telephone service interruption was significant for only two vaccinations, 3+Hib and 3+HepB. Geographic variables, when significant, were generally among the last variables to enter the model.

Child's race/ethnicity and age and household size were significant for all vaccinations, and typically were one of the first five variables to enter the model. Mother's age, education, and marital status as well as insurance coverage and number of children were significant for a large majority of vaccinations, although only insurance coverage was typically one of the first five variables to enter the model.

4. Implications for NIS Ratio Adjustment

The Keeter adjustment is applied based on the assumption that households without landline are similar to the households with interruption in telephone service. This assumption is based on studies conducted when the majority of non-landline households were phoneless; however, with rapidly increasing use of cell phone the majority of the non-landline households are now cell-only users.

Srinath, et al. (2009) evaluate two alternatives to the Keeter approach with 2001-2002 NHIS data – simple post-stratification to NHIS age, gender, and race/ethnicity categories and a propensity adjustment using a logistic regression to assess the probability of a households being non-telephone. While their results showed the Keeter approach yielded the smallest MSE, there are limitations associated with their overall approach given the current state of telephony.

---

[3] 4 or more doses of DTaP.
[4] 3 or more doses of poliovirus vaccine.
[5] 1 or more doses of any MMR.
[6] 3 or more doses of Hib.
[7] 3 or more doses of HepB.
[8] 1 or more doses of varicella vaccine.
[9] 3 or more doses of pneumococcal conjugate vaccine (PCV).
[10] 4 or more doses of DTaP, 3 or more doses of poliovirus vaccine, 1 or more doses of any MMR, 3 or more doses of Hib, 3 or more doses of HepB, 1 or more doses of varicella vaccine.
[11] 4:3:1 plus 3 or more doses of Hib.
[12] 4:3:1:3 plus 3 or more doses of HepB.

First and foremost, Srinath et al. were solely looking at adjusting weights for telephone households (including cell telephone households) to account for non-telephone households. In addition, the prevalence of non-landline households was much lower in the time period they were examining (less than 3%) than encountered today (above 22%). Thus, they were attempting to adjust for a smaller undercoverage than encountered in an RDD survey such as the NIS.

In addition, the simple post-stratification did not account for all the variables highly correlated with telephone status, e.g., it did not utilize household income and mother's age and education level. Thus, one would not expect the simple post-stratification to do well in adjusting for non-telephone households. The propensity adjustment did identify the expected variables associated with telephone status (including income, education level, and health insurance coverage status).

Chowdhury, et al. (2008) utilized data from the 2006-2007 NIS and the 2006 NHIS to examine the Keeter's adjustment step within NIS weighting. The authors concluded that, while differences in child characteristics do exist between households with and without service interruptions, the impact of the Keeter's adjustment was not significant at either the national or estimation area level, although they did express concern that such results could change should the proportion of non-landline households increase.

Chowdhury, et al. also concluded that children in households with interruption are more similar to those in cell-only household than to those in phoneless households, which would have potential implications for application of and potential modification of.the Keeter's adjustment step. However, the results of this research suggest that the more important issue is controlling for correlated variables than attempting to align by telephone status and/or telephone service interruption.

4.1 Possible Alternative Approaches

Results obtained in the current analysis suggest promise for improved accuracy from several alternate approaches. All involve eliminating the Keeter adjustment and combining the household level post-stratification and raking adjustments.

First, a raking adjustment could be applied incorporating geographic and sociodemographic variables correlated with vaccination status. The results of this analysis also suggest determination of the appropriate sociodemographic variables could be made using NIS data alone, as it does not appear that telephone characteristics are important predictors. One additional sociodemographic variable, owner/renter status, has become available effective Q3/2008 for inclusion in the determination.

A second approach would be to use weight adjustments based on model-based propensities of having non-landline telephone, as suggested by Khare, *et al*. (2009). Using information on sociodemographics, geography, access to landline and wireless phones, and telephone service interruption, propensities could be modeled using the NHIS data, with the model information then applied to the NIS data to form weighting classes based on the propensity quintiles and other geography and sociodemographic covariates. To assess the potential for bias, NIS vaccination coverage estimates could be compared across quintiles of predicted propensities associated with having no access to landline telephones.

A third approach would be to replace the ratio adjustments with a weighting methodology yielding weights consistent with derivation of a GREG estimator,

$$\hat{Y}_{vG} = \hat{Y}_{vS} + \hat{\mathbf{B}}'\left(\mathbf{T}_x - \hat{\mathbf{T}}_{xS}\right)$$

where $\hat{Y}_{vS}, \hat{Y}_{vG}$ represent the estimates pre-and post-GREG for vaccination variable v, $\mathbf{T}_x$ represents the vector of population totals for the auxiliaries, $\hat{\mathbf{T}}_{xS}$ represents the estimate of $\mathbf{T}_x$ prior to GREG, and $\hat{\mathbf{B}}'$ represents the slope of the GREG. Significant predictors of vaccination status would be used in the vector of covariates. One advantage of applying a GREG estimator approach would be greater control over the change in weights resulting from the application of population controls. Hedlin, et al. (2001) have pointed out that the performance of GREG estimators can be sensitive to model misspecification, although the example they state was in the situation of an establishment survey, which tend to have skewed population distributions.

5. Conclusion

The analysis presented in this paper suggests that, while the NIS must adjust for undercoverage of non-landline telephone households, it is not necessary to use population controls based on telephone characteristics. Rather, use of appropriate sociodemographic variables correlated with variables of interest may be sufficient. While it is likely the sociodemographic variables selected will also be correlated with telephone status, that is not a requirement nor is it necessary to utilize all variables correlated with telephone status – only those correlated with the variables of interest.

The characteristics of the children in the households with no landline telephone and those with a landline telephone do appear to differ on race/ethnicity, income, and mother's age and education. These variables also appear to be correlated with some health measures of interest for the NIS.

Next steps in considering a refinement of the NIS weighting methodology should include:

1) determination of the full set of sociodemographic variables to include in adjustments;
2) derivation of model-based propensities from NHIS data;
2) reweighting of a recent set of NIS data using the single raking step, propensity approach, and GREG, applying the new set of variables in the raking and GREG;
3) comparison of weight distributions and variance estimates between the current and alternative methodologies

Finally, should the refined methodology be determined to offer improved accuracy for NIS estimates, consideration will have to be given as to the appropriate manner in which such a refinement could be implemented to support use of NIS estimates in examining trends over time.

References

Blumberg SJ, Luke JV. *Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2008*. National Center for Health Statistics. May 2009. Available from: http://www.cdc.gov/nchs/nhis.htm

CDC (2008). *2007 National Immunization Survey Methodology Report*. CDC, Atlanta.

Chowdhury, S.R., Montgomery, R., and Smith, P.J. (2008). *Adjustment for Noncoverage of Non-landline Telephone Households in an RDD Surveys*. 2008 Proceedings of the Section on Survey Research Methods [CD-ROM], Alexandria, VA: American Statistical Association: pp 1438-1443.

Frankel, M.R., Srinath, K.P., Hoaglin, D.C., Battaglia, M.P., Smith, P.J., Wright, R.A., and Khare, M. (2003). *Adjustments for non-telephone bias in random-digit-dialing surveys.* Statistics in Medicine, Volume 22, pp. 1611-1626.

Hedlin, D., Falvey, H., Chamber, R., and Kokic, P. (2001). *Does the ModelMatter for GREG Estimation? A Business Survey Example*. Journal of Official Statistics, Volume 17, pp. 527-544.

Keeter S. (1995) *Estimating non-coverage bias from a phone survey*. Public Opinion Quarterly; 59:196–217.

Khare, M., Chowdhury, S., Wolter, K., and Singleton, J.A. (2007). *An Assessment of Bias due to Noncoverage of Wireless Only Households in RDD Surveys*. 2007 Proceedings of the Section on Survey Research Methods [CD-ROM], Alexandria, VA: American Statistical Association: pp 3911-3918.

Khare, M., Wouhib, A., and Singleton, J.A. (2009). Assessment of Potential Bias in the National Immunization Survey Using Data on Telephone Status from the 2007 NHIS. Presented at the 2009 Conference of the American Association of Public Opinion Research.

Molinari, N.M., Singleton, J., Khare, M., Smith, P., Wolter, K., Santos, K., Montgomery, R., Chowdhury, S., Barron, M., Skalland, B., Copeland, K. (2008). *The distribution of total error in a health survey: a progress update***.** Presented at the International Total Survey Error Workshops 2008, RTI, North Carolina.

National Health Interview Survey (NHIS). ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2005/srvydesc.pdf, 2005

Smith, P.J., Hoaglin, D.C., Battaglia, M.P., Khare, M., and Barker, L.E. (2005). *Statistical Methodology of the National Immunization Survey: 1994-2002*. National Center for Health Statistics. Vital Health Stat 2(138).

Srinath, K.P., Frankel, M.R., Hoaglin, D.C., and Battaglia, M.P. (2009). *Compensating for Noncoverage of Nontelephone Households in Random-Digit-Dialing Surveys: A Comparison of Adjustments Based on Propensity Scores and Interruptions in Telephone Service*. Journal of Official Statistics, Volume 25, pp. 77-98.

Table 1: Range of Quarterly and Annual Weighting Adjustments by Weighting Stage, 2007 NIS

Weighting Stage	Weighting Stage Adjustment			Cumulative Adjustment		
	5th Pct	Median	95th Pct	5th Pct	Median	95th Pct
Quarterly Level Weighting Adjustments						
Resolution NR Adjustment	1.1	1.3-1.4	1.4-1.5	1.1	1.3-1.4	1.4-1.5
Screener NR Adjustment	1.1	1.1	1.2	1.3	1.4-1.5	1.6-1.7
Interview NR Adjustment	1.0-1.1	1.1-1.2	1.3	1.3-1.4	1.6-1.7	1.9-2.1
Annual Level Weighting Adjustments						
Keeter Adjustment	**1.6**	**2.1**	**3.6**	1.6	2.1	3.6
Post-stratification Adjustment	0.6	0.8	1.8	1.1	1.8	4.6
Raking Adjustment	0.7	1.0	1.3	1.0	1.6	5.0
Provider NR Adjustment	1.2	1.4	1.9	1.3	2.4	7.6
Raking Adjustment	0.8	1.0	1.2	**1.3**	**2.4**	**7.8**

Table 2: Weighted Sample Distributions from the 2007 NIS and the 2007 NHIS for Children Aged 19-35 Months, by Selected Characteristics

	NIS Distribution	NHIS Distribution	Difference (NIS-NHIS)
Race/Ethnicity			
Hispanic	27.5	24.0	3.5
NH-White	51.2	55.1	-3.9
NH-Black	12.5	15.0	-2.5
NH-Other	8.7	5.8	2.9
Ratio of Income to Poverty			
< 1.0	29.6	18.2	11.3
1.0 - 1.99	22.7	18.2	4.5
2.0 - 3.99	24.7	21.6	3.1
4.0+	23.1	16.0	7.1
Missing/Dk*		26.0	
Household Size			
2	2.1	4.5	-2.4
3 or 4	53.8	58.7	-5.0
5+	44.1	36.8	7.4
Mother's Education			
< 12 years	20.5	19.7	0.8
12 years	30.5	19.6	10.9
> 12 years	49.0	55.8	-6.8
Missing/Dk*		4.9	
Mother's Age			
< 30	43.0	48.1	-5.1
30+	57.0	47.4	9.6
Missing/Dk*		4.5	

* Missing/Don't Know (DK) values imputed as part of NIS data processing

Table 3: Weighted Prevalences and Sample Distributions from the 2007 NHIS for Children Aged 19-35 Months, by Telephone Status

	With Landline Telephone	With No Landline Telephone			Difference (Landline - No Landline)
		Total	Cell Phone	No Phone	
Prevalence	73.5	26.5	22.1	4.4	
Weighted Sample Distributions by Selected Characteristics					
Race/Ethnicity					
Hispanic	**21.7**	**30.3**	31.3	25.2	-8.6
NH-White	**58.5**	**45.9**	46.6	42.4	12.5
NH-Black	**14.2**	**6.4**	19.8	5.1	7.8
NH-Other	**5.6**	**6.4**	2.3	27.3	-0.9
Ratio of Income to Poverty					
< 1.0	**13.0**	**32.7**	35.7	17.9	-19.7
1.0 - 1.99	**17.4**	**20.3**	22.0	11.4	-2.8
2.0 - 3.99	**23.9**	**15.1**	17.1	4.8	8.8
4.0+	**19.8**	**5.3**	5.5	4.7	14.5
Missing/Dk	**25.8**	**26.6**	19.7	61.2	-0.8
Household Size					
2	**3.3**	**8.0**	8.6	4.7	-4.7
3 or 4	**59.9**	**55.5**	59.6	34.9	4.4
5+	**36.8**	**36.5**	31.8	60.4	0.3
Mother's Education					
< 12 years	**15.0**	**32.7**	28.2	55.4	-17.7
12 years	**19.1**	**21.0**	23.5	8.6	-2.0
> 12 years	**61.7**	**39.4**	43.2	19.8	22.4
Missing/Dk	**4.2**	**6.9**	5.0	16.2	-2.7
Mother's Age					
< 30	**39.5**	**71.7**	79.3	33.4	-32.2
30+	**56.7**	**21.5**	15.8	50.4	35.2
Missing/Dk	**3.7**	**6.7**	4.9	16.2	-3.0

Table 4: Weighted Prevalences and Sample Distributions from the 2007 NIS and the 2007 NHIS for Children Aged 19-35 Months, by Telephone Status and Telephone Service Interruption

	NIS Distribution			NHIS Distribution (Final Weighted)									
	Final Weighted			Total	With Landline Telephone						With No Landline Telephone		
	Total	Service Interruption			Total	Service Interruption		Cell Phone Status			Total	Cell Phone	No Phone
		No Int	Int			No Int	Int	No Cell	Some Cell	Mostly Cell			
Prevalence				100.0	73.5			15.9	43.2	14.3	26.5	22.1	4.4
Weighted Sample Distribution for Selected Characteristics													
Race/Ethnicity													
Hispanic	**27.5**	26.4	34.4	**24.0**	**21.7**	21.5	26.7	33.8	14.6	29.8	**30.3**	31.3	25.2
NH-White	**51.2**	53.5	37.0	**55.1**	**58.5**	59.5	32.2	45.6	68.1	43.9	**45.9**	46.6	42.4
NH-Black	**12.5**	11.3	20.5	**15.0**	**14.2**	13.5	31.9	17.8	12.0	17.0	**17.3**	19.8	5.1
NH-Other	**8.7**	8.8	8.2	**5.8**	**5.6**	5.4	9.3	2.8	5.3	9.4	**6.4**	2.3	27.3
Ratio of Income to Poverty													
< 1.0	**29.6**	25.6	54.2	**18.2**	**13.0**	12.0	36.7	22.5	9.4	13.2	**32.7**	35.7	17.9
1.0 - 1.99	**22.7**	22.1	25.9	**18.2**	**17.4**	16.9	29.6	24.2	14.5	18.8	**20.3**	22.0	11.4
2.0 - 3.99	**24.7**	26.5	13.0	**21.6**	**23.9**	24.3	13.1	16.3	28.5	18.5	**15.1**	17.1	4.8
4.0+	**23.1**	25.7	6.8	**16.0**	**19.8**	20.2	9.9	5.2	23.5	24.9	**5.3**	5.5	4.7
Missing/Dk*				**26.0**	**25.8**	26.4	10.8	31.7	24.0	24.6	**26.6**	19.7	61.2
Household Size													
2	**2.1**	1.9	3.7	**4.5**	**3.3**	3.2	5.5	3.0	2.3	6.5	**8.0**	8.6	4.7
3 or 4	**53.8**	54.3	50.7	**58.7**	**59.9**	60.1	55.5	56.5	60.4	62.2	**55.5**	59.6	34.9
5+	**44.1**	43.9	45.6	**36.8**	**36.8**	36.7	39.0	40.5	37.3	31.4	**36.5**	31.8	60.4
Mother's Education													
< 12 years	**20.5**	18.2	35.1	**19.7**	**15.0**	14.8	20.7	25.6	11.8	13.0	**32.7**	28.2	55.4
12 years	**30.5**	30.0	33.5	**19.6**	**19.1**	19.1	19.1	28.8	15.4	19.3	**21.0**	23.5	8.6
> 12 years	**49.0**	51.8	31.4	**55.8**	**61.7**	62.2	50.1	38.7	69.9	62.6	**39.4**	43.2	19.8
Missing/Dk*				**4.9**	**4.2**	4.0	10.1	6.9	2.9	5.1	**6.9**	5.0	16.2
Mother's Age													
< 30	**43.0**	40.1	61.0	**48.1**	**39.5**	38.8	58.6	44.8	31.6	57.7	**71.7**	79.3	33.4
30+	**57.0**	59.9	39.0	**47.4**	**56.7**	57.8	31.3	48.9	66.0	37.6	**21.5**	15.8	50.4
Missing/Dk*				**4.5**	**3.7**	3.5	10.1	6.4	2.4	4.8	**6.7**	4.9	16.2

* Missing/Don't Know (DK) values imputed as part of NIS data processing

Table 5: Logistic Regression Modeling Results for Selected Health Status Variables, 2007 NHIS

Variables Common to NIS, NHIS	Flu Shot p-value		Asthma p-value		Chicken Pox p-value	
	Full Model[1]	Final Model[2]	Full Model[1]	Final Model[2]	Full Model[1]	Final Model[2]
Geography						
Region	0.003	0.0004	0.001	0.007	0.008	<0.0001
MSA Status	0.056	0.043	0.530		0.369	
Child Demographics						
Race/Ethnicity	0.052	0.028	<0.0001	<0.0001	0.056	0.035
Gender	0.558		0.005	0.004	0.373	
Age	0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Mother Demographics						
Age	0.805		0.022	0.009	0.014	0.006
Education	0.510		0.005	0.026	0.463	
HH Characteristics						
Income to Poverty Ratio	0.196	0.011	0.059	0.009	0.001	0.002
Insurance Coverage	0.010	0.003	0.012	0.015	0.135	0.047
Size	0.801		0.494		0.495	
Telephone Status						
LL/Cell/None	0.095		0.206		0.835	

[1] Prior to stepwise logistic regression
[2] Following stepwise logistic regression

Table 6: Logistic Regression Modeling Results for Selected Health Status Variables, 2007 NIS

Variables Common to NIS, NHIS	Flu Shot p-value		4:3:1:3:3:1* p-value	
	Full Model[1]	Final Model[2]	Full Model[1]	Final Model[2]
Geography				
Region	0.493		0.085	
MSA Status	0.053		0.593	
Child Demographics				
Race/Ethnicity	0.523		0.144	
Gender	0.958		0.699	
Age	<0.0001	<0.0001	<0.0001	<0.0001
Mother Demographics				
Age	0.956		0.093	0.010
Education	0.096	0.040	0.715	
HH Characteristics				
Income to Poverty Ratio	0.005	0.0002	0.203	
Insurance Coverage	0.823		0.002	0.001
VFC	0.281		<0.0001	<0.0001
Size	0.466		0.0002	<0.0001
Telephone Service				
Interruption Status	0.762		0.842	

* 4 or more doses of DTaP, 3 or more doses of poliovirus vaccine,
1 or more doses of any MMR, 3 or more doses of Hib,
3 or more doses of HepB, 1 or more doses of varicella vaccine

[1] Prior to stepwise logistic regression
[2] Following stepwise logistic regression

Table 7: Logistic Regression Modeling Results for Vaccination Series Using Full Set of NIS Variables, 2007 NIS

Variable	Vaccination											Summary Information	
	4+DTaP[‡]	3+Polio[§]	1+MMR[II]	3+Hib[¶]	3+HepB**	1+Var[††]	3+PCV[‡‡]	4:3:1[IIII]	4:3:1:3[¶¶]	4:3:1:3:3***	4:3:1:3:3:1[†††]	# of Vaccinations	Median Order[1]
	Order of Entry into Logistic Regression Model												
Geography													
Region	11					7	10	10		9	7	6	10
MSA Status	12		10			6	2					4	12
Child Demographics													
Race/Ethnicity	5	3	3	6	4	4	4	4	4	4	4	11	4
Gender	10		5					11				3	12
Age	1	2	1	1	2	1	11	1	1	1	1	11	1
First-born Status			7		7			9	7	8	8	6	9
Mother Demographics													
Age	9		11	3			9	8	6	7	6	8	9
Education	7	4			6	8	3	5	8	5	9	9	7
Marital Status	6		6	5			8	3	3	6	5	8	6
HH Characteristics													
Income to Poverty Ratio	2		8	9	5	9	7					6	9
Insurance Coverage	4		4	4		2	6	6	5	3	3	9	4
Size	3	1	2	2	1	3	1	2	2	2	2	11	2
Number of Children	8	5	9	7		5	5	7	9	10	10	10	8
Telephone Service													
Interruption Status				8	3							2	12.5

[‡] 4 or more doses of any diphtheria and tetanus toxoids and pertussis vaccines including diphtheria and tetanus toxoids, and any acellular pertussis vaccine (DTaP/DTP/DT).

[§] 3 or more doses of any poliovirus vaccine.

[II] 1 or more doses of measles-mumps-rubella vaccine.

[¶] 3 or more doses of *Haemophilus influenzae* type b (Hib) vaccine.

** 3 or more doses of hepatitis B vaccine.

[††] 1 or more doses of varicella at or after child's first birthday, unadjusted for history of varicella illness.

[‡‡] 3 or more doses of pneumococcal conjugate vaccine (PCV).

[IIII] 4 or more doses of DTaP, 3 or more doses of poliovirus vaccine, and 1 or more doses of any MMR.

[¶¶] 4:3:1 plus 3 or more doses of Hib.

*** 4:3:1:3 plus 3 or more doses of HepB.

[†††] 4:3:1:3:3 plus 1 or more doses of varicella vaccine.

[1] Order for variables not deemed significant imputed as average of places not taken by significant variables