Workshop on Using the National Survey of Children’s Health Dataset: Practical Applications

Julian Luke
Stephen Blumberg

Centers for Disease Control and Prevention
National Center for Health Statistics

SAFER • HEALTHIER • PEOPLE™
National Survey of Children’s Health, 2003

- Sponsored by HRSA’s Maternal and Child Health Bureau
- Conducted by CDC’s National Center for Health Statistics
- Part of the State and Local Area Integrated Telephone Survey (SLAITS) program
Agenda

- Brief description of the NSCH
- Key sociodemographic variables
- Weights and sampling variables
- Tips on preparing data for analysis
- Frequencies, crosstabs, and logistic regression examples in SUDAAN, SAS, and STATA
To produce national and state-based estimates on the health and well-being of children, their families, and their communities.
Interview Process

- Independent random-digit-dial samples for all 50 states plus D.C.
- Screened households for children under 18 years of age
- One child under 18 years of age was randomly selected to be the target of the interview
Final Sample

- 102,353 completed interviews
  - Minimum: 1,483 in Utah
  - Maximum: 2,241 in Louisiana and Ohio
  - 25 states have more than 2,000 each

- Overall response rate: 55.3%
  - Minimum: 49.6% in New Jersey
  - Maximum: 64.4% in South Dakota
  - 32 states were above 55%
Weighting and Estimation

- Sampling weights to permit national and state-specific estimates of health and well-being
- Sampling weights are adjusted for potential non-response biases
- Sampling weights are adjusted to account for non-coverage of non-telephone households
Locating SLAITS Data

- The SLAITS home page is located at http://www.cdc.gov/nchs/slaits.htm
- SLAITS Overview
- Key Features of the SLAITS Mechanism
- Frequently Asked Questions
- Existing Survey Modules
Key Variables on Public Use File
Key Variables on Public Use File:

State of Residence

- Name: STATE
- Levels: Separate numeric code for each state
Key Variables on Public Use File:

**Age**

- **Name:** AGEYR_CHILD
- **Levels:** Age in years
  - 0 = Younger than one year
Key Variables on Public Use File:

**Sex**

- Name: S1Q01
- Levels: 1 = male, 2 = female
Confidentiality

- Confidentiality was guaranteed to participants
- Section 308d of the Public Health Service Act (42 U.S.C. 242m):
  
  “No information…may be used for any purpose other than the purpose for which it was supplied…[and] may not be published or released…if the particular establishment or person supplying the information or described in it is identifiable.”

- Prohibits the release of sub-state identifiers or contextual information
Key Variables on Public Use File:

**Race**

- **Name:** RACER, RACEAIAN, RACEASIA, RACE_HI
- **Levels:**
  - RACER = White, Black, Other, Multirace
  - RACEAIAN adds American Indian/AK Native
  - RACEASIA adds Asian
  - RACE_HI adds Asian and Native Hawaiian / PI
- **Only RACER can be used for national estimates**
Key Variables on Public Use File:

Family Structure

- Name: FAMSTRUCT
- Levels:
  - 1 = Two-parent biological/adoptive household
  - 2 = Two-parent household with at least one step-parent
  - 3 = One-parent household with a biological, step, foster, or adoptive mother and no father
  - 4 = All other family structures
Key Variables on Public Use File:

Relative Ages of Children

Name: AGEPOS4

Levels:

- 1 = Only child
- 2 = Oldest child
- 3 = Second oldest child
- 4 = Third oldest child
- 5 = Fourth oldest child or younger

Note: This variable refers to the relative ages of children in the household. If the child has siblings over 17 years of age or unrelated children live in the household, this variable should not be interpreted as birth order.
Key Variables on Public Use File:

**Urban/Rural Identifier**

- **Name:** MSA_STAT
- **Levels:** 1 = Yes, 0 = No
- **Missing Data:** MSA_STAT was suppressed in 16 states to protect the confidentiality of participants
Key Variables on Public Use File:

**Income**

- **Name:** POVERTY_LEVELR
- **Levels:** 8 categories relative to the Federal Poverty Level
- **Derived from:** Total number of household members and household income value
- **Missing Data:** Total household members and/or household income were missing
Key Variables on Public Use File:

**Highest Education Achieved**

- **Name:** EDUCATIONR
- **Levels:**
  - 1 = < 12 years
  - 2 = 12 years (including high school graduate)
  - 3 = More than high school (including college graduates)
Key Variables on Public Use File:

Primary Language in Home

- Name: PLANGUAGE
- Levels:
  - 1 = English
  - 2 = Any other language
Key Variables on Public Use File:

**Number of Adults in HH**

- **Name:** TOTADULT3
- **Note:** This variable refers to the number of adults in the household. It has been topcoded at 3+. 
Key Variables on Public Use File:

**Number of Children in HH**

- **Name:** TOTKIDS4
- **Note:** This variable refers to the number of children in the household. It has been topcoded at 4+. 
Top-Coded and Bottom-Coded Variables

- How many times (past 12 months) sample child (S.C.) saw a doctor, nurse, or other health care professional for preventative medical care? (S4Q03R)
- How many times (past 12 months) S.C. go to a hospital emergency room about health? (S4Q04R)
- How many emergency room visits because of accident, injury, or poisoning (S4Q05R)
- Excluding emergency room visits, hospitalizations, and well-child care, how many times in last 12 months did S.C. see a doctor, nurse, or other health care professional for sick-child care? (S4Q06R)
Top-Coded and Bottom-Coded Variables

- How old was S.C. when completely stopped breastfeeding or being fed breast milk? (S6Q60R)
- During past 12 months, about how many days did S.C. miss school because of illness or injury? (S7Q02R)
- During the past week, how many times did you or a family member take S.C. on any kind of outing (park, library, zoo, shopping, church, etc)? (S8Q01R)
- About how often does S.C. attend a religious service? (S8Q02R)
- How many times has S.C ever moved to a new address? (S11Q06R)
Key Variables on Public Use File:

**Overweight**

- **Name:** BMICLASS
- **Levels:**
  - 1 = Underweight
  - 2 = Normal weight
  - 3 = At risk of overweight
  - 4 = Overweight
- **Derived from:** Parent-reported height and weight, which are top-coded and bottom coded (see HGHT_FLG and WGHT_FLG)
Weights and Sampling Variables
Use Weighted Data

- Sampling weights to permit national and state-specific estimates
- Sampling weights are adjusted for potential non-response biases
- Sampling weights are adjusted to account for non-coverage of non-telephone households
Only One Weight

- WEIGHT_I

→ The same weight is used for national and state-level analyses.
Variance Estimation

- Sample design involved clustering of children within households and stratification of household within states.
- Therefore, SUDAAN, SAS survey procedures, STATA, or other such programs must be used to obtain estimates of variability and statistical significance.
Several data users have noted that, within a given state, the NSCH sample was drawn as a simple random sample. Therefore, these data users have suggested that accurate variance estimates for a single state can be obtained from SAS if the sampling weights are normalized.

This is not true. SAS does not adequately account for the heterogeneous sampling weights. Therefore, the standard errors provided by SAS for normalized weights are smaller than they should be.
Sampling Variables

- Stratum: State (STATE)
- PSU: Household (IDNUMR)

In SUDAAN…
- PROC ... DESIGN=WR;
- NEST STATE IDNUMR;
- WEIGHT WEIGHT_I;
Statistical packages used to analyze Survey Data

- SUDAAN – SUrvey DAta ANalysis
- Selected SAS procedures: SURVEYFREQ and SURVEYLOGISTIC
- STATA

- Taylor series linearization method for estimating population characteristics from complex survey data
Preparing the Data for Analysis

- Subsetting the population to analyze particular subgroups only
- Dealing with question non-response values like “Don’t Know” and “Refused” responses
- Categorical variables in SUDAAN
Preparing the Data: Subsetting

- Don’t subset (e.g., SAS “subsetting if statement”, deleting unneeded records)
- Subsetting can delete entire PSUs from the sample design
- The software needs to “see” the entire design structure to accurately estimate the standard errors
- Use specific procedural statements (e.g., SUPPOPN) to specify an analytic subgroup (e.g., males)
Preparing the Data: Item Non-Response

- Set “Don’t Know” and “Refused” responses to missing values
  - DK = 6, 96, 996, 9996
  - RF = 7, 97, 997, 9997
- Variables with missing values set to “.” are excluded from the analysis
Preparing the data:

Categorical Variables

- SUDAAN variables require special preparation for some of types of analysis
  - Categorical variables must begin with 1 and increase in whole numbers with no integers missing
  - Yes/no – 1/0 recode to 1/2
  - Subgroup and Levels statements
  - Will become clear when we look at sample programs
Statistics

- Frequency distributions and Crosstabs
- Logistic regression
General SAS and SUDAAN Syntax

**PROC _____ ;**

- Sample Design statements
- Computational statements
- Output specifications
PROC CROSSTAB DESIGN = WR;
  NEST STATE IDNUMR;
  WEIGHT WEIGHT_I;
  SUBPOPON VAR1 = value;
  SUBGROUP VAR2 VAR3 VAR4;
  LEVELS n2 n3 n4;
  TABLES VAR2 VAR3*VAR4;
  PRINT NSUM WSUM ROWPER SEROW
   / STYLE = NCHS WSUMFMT = F7.0
   SEROWFMT = F5.4;
*** Create Analytical Variables ***
*** To use PROC CROSSTAB SUITAN requires variables not to have a zero value ***
*** All DF-No values must be coded to 0/Non***

/* Use Special Needs Screener to Classify a Child as Having a Special Need */
cehcn=sq;
if sq06=1 or sq05=1 or sq12=1 or sq18=1 or sq17=1 then cehcn=1;

/* sq06 - Prescription Meds for a condition of 12 months or longer */
/* sq05 - Medicaid care, mental health, ed services for 12 months or longer */
/* sq12 - Limitation in abilities for a condition for 12 months or longer */
/* sq18 - Need special therapy for a condition for 12 months or longer */
/* sq17 - Emotional, behav problem for a condition for 12 months or longer */

health = .;
if sq01 in (1,2) then health=1; /* Health Status Excellent/Very Good */
if sq01 in (3,4,5) then health=2; /* Health Status Good/Fair/Poor */

health=health;
if health=2 then health=0;

age41 = .;
if 0 <= ageyrs_child <= 6 then age41 = 1;
if 6 <= ageyrs_child <= 11 then age41 = 2;
if 12 <= ageyrs_child <= 17 then age41 = 3;

/* sq03 - Male or Female */
/* sq02 - Ever told by doctor had asthma */
/* sq03 - Ever told by doctor had hearing or vision problems */
/* sq02 - Ever told by doctor had ADD */
/* racer - Race classification */
/* sq08r - How many times gone to doctor for sick care */

*** Recode Values to Missing ***
array miss(*) sq01 sq03 sq09 sq020 sq021 racer;
do 1 = 1 to dim(miss);
   if miss(1) in (6 7 L .M H) then miss(1) = .;
end;
if sq08r in (96 97 .L .M .H) then sq08r = .;

*** Recode 1-No to 2-No ***
array ano(*) sq019 sq020 sq021;
do 1 = 1 to dim(ano);
   if ano(1) = 0 then ano(1) = 2;
end;

*** Create New Variable Labels ***
label
  cehcn  = 'Child has Special Health Care Needs'
  health = 'Health Status'
  age41  = 'Age at Study'
  miss(1) = 'Missing'
  ano(1)  = 'Has Some Kind of Illness For Sick Care'
proc forest;
value yesno
1='Yes'
2='No' ;
value age19
1 = '0-5 Years'
2 = '6-11 Years'
3 = '12-17 Years' ;

/* SUDAAN requires sorting by STRATA and PSU prior to SURVEY PROCES */
proc sort;
by state idnum;
run;

/* Demonstrate the CROSSTAB Procedure */
title2 'Univariate Frequencies on Various Variables, All Children';
proc crosstab design='r';
  nest state idnum;
  weight weight_1;
  subgroup s1q01 race+ agel9 s2q19 s2q20 s2q21 health cshcn;
  levels 2 4 2 2 2 2 1 ;
  table s1q01 race+ agel9 s2q19 s2q20 s2q21 health cshcn ;
  rformat agel9 5.2g ;
  rformat cshcn yesno ;
  rformat health yesno ;
  rformat s2q19 yesno ;
  rformat s2q20 yesno ;
  rformat s2q21 yesno ;
  set readwidth=14 decwidth=4;
  print num wsm  rscore rscore / style=XCHS;
run;

title1 "Subpopulation Analysis";
title2 'Multi-variate Frequ: Overall Health by Asthma Status';
title3 "Using SUBPOP Statement - Where RACE = White";
proc crosstab design='r';
  nest state idnum;
  weight weight_1;
  subgroup race=1;
  subgroup race 0 s2q19 health ;
  levels 4 2 2 ;
  table s2q19 * health ;
  rformat health yesno ;
  rformat s2q19 yesno ;
  set readwidth=14 decwidth=4;
  print num wsm  rscore / style=XCHS;
run;
The SAS System  
Univariate Frequencies on Various Variables, All Children

Date: 12-02-2000  
Time: 12:46:54  
Research Triangle Institute  
The CRORSSTAB Procedure  
Page : 1  
Table : 1

Variance Estimation Method: Taylor Series (VR)  
by: [S.C.] male or female?

<table>
<thead>
<tr>
<th></th>
<th>Sample Size</th>
<th>Weighted Size</th>
<th>Row Percent</th>
<th>SE Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>103979.0000</td>
<td>72693628.8780</td>
<td>100.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1 - MALE</td>
<td>52264.0000</td>
<td>37782066.8180</td>
<td>51.1200</td>
<td>0.2734</td>
</tr>
<tr>
<td>2 - FEMALE</td>
<td>51715.0000</td>
<td>34911562.0604</td>
<td>48.8797</td>
<td>0.2734</td>
</tr>
</tbody>
</table>

Variance Estimation Method: Taylor Series (VR)  
by: Race classification for all states [White,Black,Multiracial,Other].

<table>
<thead>
<tr>
<th>Race classification for all states [White,Black,Multiracial,Other]</th>
<th>Sample Size</th>
<th>Weighted Size</th>
<th>Row Percent</th>
<th>SE Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>95441.0000</td>
<td>68482956.5090</td>
<td>100.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1 - WHITE ONLY</td>
<td>75408.0000</td>
<td>50830511.6200</td>
<td>77.4618</td>
<td>0.2711</td>
</tr>
<tr>
<td>2 - BLACK ONLY</td>
<td>10134.0000</td>
<td>7071195.6604</td>
<td>7.4049</td>
<td>0.2244</td>
</tr>
<tr>
<td>3 - MULTIRACE</td>
<td>4407.0000</td>
<td>3203867.5162</td>
<td>3.3718</td>
<td>0.1041</td>
</tr>
<tr>
<td>4 - OTHER</td>
<td>4499.0000</td>
<td>3395888.4771</td>
<td>3.5874</td>
<td>0.1802</td>
</tr>
</tbody>
</table>

Variance Estimation Method: Taylor Series (VR)  
by: Has a doctor or health professional ever told you that [S.C.] has asthma?

<table>
<thead>
<tr>
<th>Has a doctor or health professional ever told you that [S.C.] has asthma?</th>
<th>Sample Size</th>
<th>Weighted Size</th>
<th>Row Percent</th>
<th>SE Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>102185.0000</td>
<td>72506181.9270</td>
<td>100.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Yes</td>
<td>12502.0000</td>
<td>9045296.2299</td>
<td>12.2456</td>
<td>0.1089</td>
</tr>
<tr>
<td>No</td>
<td>89683.0000</td>
<td>63060885.7040</td>
<td>87.7544</td>
<td>0.1868</td>
</tr>
</tbody>
</table>
Subpopulation Analysis

Multi-variant Frege: Overall Health by Asthma Status
Using SUBPOP Statement: WHERE RACE = White

12:44 Saturday, December 2, 2006

SUDAAN
Software for the Statistical Analysis of Correlated Data
Copyright Research Triangle Institute February 2006
Release 9.0.1

Number of observations read: 102353
Weighted count: 72796925
Observations in subpopulation: 76408
Weighted count: 488807192
Bonferroni degrees of freedom: 70208

Date: 12-02-2006
Time: 12:47:00
Research Triangle Institute
The CROSSTAB Procedure
Page: 1
Table: 1

Variance Estimation Method: Taylor Series (UR)
For Subpopulation: RACE = 1
By: Has a doctor or health professional ever told you that [S.C.] has asthma? Child Overall Health is Excellent/Very Good.

--------------------------------------------------------------------------------------------------
Has a doctor or health professional ever told you that [S.C.] has asthma? Child Overall Health is Excellent/Very Good.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Weighted Size</th>
<th>Raw Percent</th>
<th>95% Raw Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>68468.0000</td>
<td>40104534.2909</td>
<td>78.3595</td>
</tr>
<tr>
<td>No</td>
<td>7771.0000</td>
<td>56692249.6120</td>
<td>11.6405</td>
</tr>
</tbody>
</table>

Yea
<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Weighted Size</th>
<th>Raw Percent</th>
<th>95% Raw Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>68468.0000</td>
<td>40104534.2909</td>
<td>78.3595</td>
</tr>
<tr>
<td>Yes</td>
<td>5510.0000</td>
<td>4071555.2139</td>
<td>73.2506</td>
</tr>
<tr>
<td>No</td>
<td>2001.0000</td>
<td>143555.0000</td>
<td>26.7494</td>
</tr>
</tbody>
</table>

No
<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Weighted Size</th>
<th>Raw Percent</th>
<th>95% Raw Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>67646.0000</td>
<td>40227632.4089</td>
<td>78.0000</td>
</tr>
<tr>
<td>Yes</td>
<td>62156.0000</td>
<td>38803839.0071</td>
<td>95.0350</td>
</tr>
<tr>
<td>No</td>
<td>5680.0000</td>
<td>4194724.6518</td>
<td>9.9650</td>
</tr>
</tbody>
</table>
Logistic Regression:
SUDAAN LOGISTIC/RLOGIST Syntax

- PROC LOGISTIC DESIGN=WR;
- PROC RLOGIST DESIGN=WR;
  - NEST STATE IDNUMR;
  - WEIGHT WEIGHT_I;
  - SUBGROUP IVAR1 IVAR2;
  - LEVELS 2 4
  - MODEL DVAR = IVAR1 IVAR2 IVAR3;
  - REFLEVEL IVAR1=Value1 IVAR2=Value2;
Logistic Regression:

SUDAAN LOGISTIC/RLOGIST Syntax

- PROC LOGISTIC DESIGN=WR;
- PROC RLOGIST DESIGN=WR;
  - NEST STATE IDNUMR ;
  - WEIGHT WEIGHT_I ;
  - MODEL DVAR = IVAR1M IVAR2H IVAR2B IVAR2O IVAR3;
data temp;
  set puf.schpuf5_formatted;

*** Create Analytical Variables ***;
/* s2q54 - How would you describe condition of child’s teeth */
/* q1o1 - Male or Female: Male = 1 */
/* rater - Race classification: white, black, multiple, other */
/* s1q01 - Is sample child of hispanic origin: Yes = 1 */

dexavg=.;
if s2q54 in (1,2) then dexavg=1; /* Condition of teeth Excellent/Very Good */
if s2q54 in (3,4) then dexavg=0; /* Condition of teeth Good, Fair, or Poor */
male=.;
if q1o1=1 then male=1; /* Male */
if q1o1=2 then male=0; /* Female */

hisrace=.;
if rater=1 then hisrace=2; /* Non-hispanic white */
if rater=2 then hisrace=9; /* Non-hispanic black */
if rater in (3,4) then hisrace=4; /* Non-hispanic other */
if s1q01=1 then hisrace=1; /* Hispanic */

hispan=.;
if hisrace=1 then hispanu=1; /* Dummy variable hispanic */
if hisrace in (2,3,4) then hispan=0;

hh_black=.;
if hisrace=3 then hh_black=1; /* Dummy variable non-hispanic black */
if hisrace in (1,2,4) then hh_black=0;

hh_other=.;
if hisrace=4 then hh_other=1; /* Dummy variable non-hispanic other */
if hisrace in (1,2,3) then hh_other=0;

*** Records Values to Missing ***;
if q1o1 in (-M,6,7) then q1o1 = .;

*** Create New Variable Labels ***;
label
  male  = "Dummy variable for males"
  hispanic = "Dummy variable for hispanics"
  hh_black = "Dummy variable for non-hispanic blacks"
  hh_other = "Dummy variable for non-hispanic others"
  hisrace = "Composite race/ethnicity variable"
  dexavg = "Child Dental Health is Excellent/Very Good"
;
run;
/* Create Value Labels */
proc format;
value hrace
1='Hispanic'
2='NH White'
3='NH Black'
4='NH Other';
run;

/* SUDAAN requires sorting by STRATA and PSU prior to SURVEY PROC */
proc sort;
by state idnum;
run;

/* Logistic Regression Using Dummy Variables */
title2 'Oral Health Ex/MG by Demographics: binary dummy variables';
proc logistic data=temp filetype=txt design=yr;
  nested idnum;
  weight weight_2;
  model dvexmg = ageyr_child mane hispanic nn_black nn_other;
run;

/* Logistic Regression Using RELEVEL Statement */
title2 'Oral Health Ex/MG by Demographics: categorical variables';
proc logistic data=temp filetype=txt design=yr;
  nested idnum;
  weight weight_2;
  subgroup s1q01 hrace;
  levels 2 4;
  model dvexmg = ageyr_child s1q01 hrace;
  relevel s1q01='2' hrace='2';
  rformat hrace.;
run;
Response variable: OVERALL: Child Dental Health as Excellent/Very Good
by: Independent Variables and Effects.

<table>
<thead>
<tr>
<th>Independent Variables and Effects</th>
<th>Beta</th>
<th>SE Beta</th>
<th>Lower 95% Limit Beta</th>
<th>Upper 95% Limit Beta</th>
<th>T-Test B=0</th>
<th>B=0</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.65</td>
<td>0.09</td>
<td>1.69</td>
<td>1.72</td>
<td>49.69</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>Derived: Age in years of selected child</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy variable for males</td>
<td>-0.04</td>
<td>0.00</td>
<td>-0.05</td>
<td>-0.04</td>
<td>-16.87</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>Dummy variable for hispanics</td>
<td>-0.09</td>
<td>0.02</td>
<td>-0.15</td>
<td>-0.04</td>
<td>-4.08</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>Dummy variable for non-Hispanic blacks</td>
<td>-1.38</td>
<td>0.04</td>
<td>-1.44</td>
<td>-1.29</td>
<td>-36.47</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>Dummy variable for non-Hispanic others</td>
<td>-0.73</td>
<td>0.04</td>
<td>-0.81</td>
<td>-0.66</td>
<td>-19.10</td>
<td>0.0008</td>
<td></td>
</tr>
</tbody>
</table>

by: Contrast.

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Degrees of Freedom</th>
<th>Wald F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVERALL MODEL</td>
<td>9</td>
<td>1014.89</td>
<td>0.0000</td>
</tr>
<tr>
<td>MODEL MINUS: INT-CERPT</td>
<td>5</td>
<td>305.55</td>
<td>0.0003</td>
</tr>
<tr>
<td>INT-CERPT</td>
<td>1</td>
<td>2408.82</td>
<td>0.0000</td>
</tr>
<tr>
<td>AGES CHILD</td>
<td>1</td>
<td>2088.12</td>
<td>0.0000</td>
</tr>
<tr>
<td>MALE</td>
<td>1</td>
<td>12.14</td>
<td>0.0001</td>
</tr>
<tr>
<td>HISPANIC</td>
<td>1</td>
<td>1380.12</td>
<td>0.0000</td>
</tr>
<tr>
<td>NH_BLACK</td>
<td>1</td>
<td>0.64</td>
<td>0.0000</td>
</tr>
<tr>
<td>NH_OTHER</td>
<td>1</td>
<td>49.59</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

by: Independent Variables and Effects.

<table>
<thead>
<tr>
<th>Independent Variables and Effects</th>
<th>Odds Ratio</th>
<th>Lower 95% Limit</th>
<th>Upper 95% Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.23</td>
<td>4.80</td>
<td>5.68</td>
</tr>
<tr>
<td>Derived: Age in years of selected child</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy variable for males</td>
<td>0.96</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>Dummy variable for hispanics</td>
<td>0.91</td>
<td>0.86</td>
<td>0.96</td>
</tr>
<tr>
<td>Dummy variable for non-Hispanic blacks</td>
<td>0.26</td>
<td>0.24</td>
<td>0.28</td>
</tr>
<tr>
<td>Dummy variable for non-Hispanic others</td>
<td>0.45</td>
<td>0.45</td>
<td>0.62</td>
</tr>
</tbody>
</table>
Response variable DVEX: Child Dental Health is Excellent/Very Good
by: Independent Variables and Effects.

<table>
<thead>
<tr>
<th>Independent Variables and Effects</th>
<th>Beta Coef.</th>
<th>SE Beta</th>
<th>Lower 95% Limit Beta</th>
<th>Upper 95% Limit Beta</th>
<th>T-Test B=0</th>
<th>T-Test B=0</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.85</td>
<td>0.20</td>
<td>1.69</td>
<td>1.72</td>
<td>4.90</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Derived Age in years of selected child</td>
<td>-0.04</td>
<td>0.00</td>
<td>-0.06</td>
<td>-0.04</td>
<td>-16.37</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Is (S.C.) male or female?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 - MALE</td>
<td>-0.02</td>
<td>0.08</td>
<td>-1.15</td>
<td>-0.04</td>
<td>-0.48</td>
<td>0.0006</td>
<td></td>
</tr>
<tr>
<td>2 - FEMALE</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Composite race ethnicity variable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-1.36</td>
<td>0.44</td>
<td>-1.44</td>
<td>-1.29</td>
<td>-6.47</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>NH White</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>NH Black</td>
<td>-0.73</td>
<td>0.44</td>
<td>-0.81</td>
<td>-0.66</td>
<td>-1.70</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>NH Other</td>
<td>-0.44</td>
<td>0.56</td>
<td>-0.56</td>
<td>-0.32</td>
<td>-0.74</td>
<td>0.0003</td>
<td></td>
</tr>
</tbody>
</table>

Response variable DVEX: Child Dental Health is Excellent/Very Good
by: Contrast.

<table>
<thead>
<tr>
<th>Contrast of Degrees of Freedom</th>
<th>Wald F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVERALL MODEL</td>
<td>1014.69</td>
<td>0.0000</td>
</tr>
<tr>
<td>MODEL MINUS INTERCEPT</td>
<td>995.58</td>
<td>0.0000</td>
</tr>
<tr>
<td>INTERCEPT</td>
<td>288.12</td>
<td>0.0000</td>
</tr>
<tr>
<td>AGE06_CHILD</td>
<td>12.14</td>
<td>0.0006</td>
</tr>
<tr>
<td>SEX01</td>
<td>400.36</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Response variable DVEX: Child Dental Health is Excellent/Very Good
by: Independent Variables and Effects.

<table>
<thead>
<tr>
<th>Independent Variables and Effects</th>
<th>Odds Ratio</th>
<th>Lower 95% Limit OR</th>
<th>Upper 95% Limit OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.23</td>
<td>4.98</td>
<td>6.60</td>
</tr>
<tr>
<td>Derived Age in years of selected child</td>
<td>0.05</td>
<td>0.02</td>
<td>0.90</td>
</tr>
<tr>
<td>Is (S.C.) male or female?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 - MALE</td>
<td>0.31</td>
<td>0.18</td>
<td>0.50</td>
</tr>
<tr>
<td>2 - FEMALE</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Composite race ethnicity variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.28</td>
<td>0.24</td>
<td>0.34</td>
</tr>
<tr>
<td>NH White</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>NH Black</td>
<td>0.48</td>
<td>0.45</td>
<td>0.62</td>
</tr>
<tr>
<td>NH Other</td>
<td>0.54</td>
<td>0.57</td>
<td>0.70</td>
</tr>
</tbody>
</table>
PROC SURVEYFREQ;
  STRATA STATE;
  CLUSTER IDNUMR;
  WEIGHT WEIGHT_I;
  TABLES VAR1 VAR2*VAR3 / display options;
  FORMAT VAR1 fmt1. VAR2 fmt2.;
*** Create Analytical Variables ***

/* Use Special Needs Screener to Classify a Child as Having a Special Need */

cshon=0;
if s2q06=1 or s2q07=1 or s2q10=1 or s2q15=1 or s2q17=1 then cshon=1;

/* s2q06 - Prescription needs for a condition of 12 months or longer */
/* s2q07 - Medical care, mental health, ad services for a condition for 12 months or longer */
/* s2q10 - Limitation in abilities for a condition for 12 months or longer */
/* s2q15 - Need special therapy for a condition for 12 months or longer */
/* s2q17 - Emotional, behavior problem for a condition for 12 months or longer */

health = :;
if s2q01 in (1,2) then health=1; /* Health Status Excellent/Very Good */
if s2q01 in (3,4,6) then health=2; /* Health Status Good/Fair/Poor */

ageg1 = :;
if 0 <= ageyr_child <= 6 then ageg1 = 1;
if 6 <= ageyr_child < 11 then ageg1 = 2;
if 12 <= ageyr_child <= 17 then ageg1 = 3;

/* s1q01 - Male or Female */
/* s2q19 - Ever told by doctor had asthma */
/* s2q20 - Ever told by doctor had hearing or vision problems */
/* s2q21 - Ever told by doctor had ADD */
/* race - Race classification */
/* v2q05 - How many lines were used to illustrate for such care */

*** Records Values to Missings ***
array write(@) s1q01 s2q19 s2q20 s2q21 race;
do i = 1 to dim(miss);
   if miss(i) in (6 7 L . M . N ) then miss(i) = ;
end;
if s2q06 in (96 97 L . M . N ) then s2q06 = ;

*** Create New Variable Labels ***
label
  cshon = "Child has Special Health Care Needs"
  health = "Child Overall Health is Excellent/Very Good"
  ageg1 = "Age of Child Grouping 1";
run;

proc format;
   value yessn
   0 = "No"
   1 = "Yes";
   value age=1g
   1 = "0-5 Years";
```sas
proc format;
  value yeano
    0 = 'No'
    1 = 'Yes'
  ;
  value age1g
    1 = '0-5 Years'
    2 = '6-11 Years'
    3 = '12-17 Years'
  ;
run;

/* SAS does not require sorting by STRATA and PSU prior to SURVEY PROCS */

/* Demonstrate the SURVEYFREQ Procedure */
title2 'Univariate Frequencies on Various Variables, All Children';
proc surveyfreq;
  strata state;
  cluster idnumr;
  weight weight_1;
  table c0q01 rcar age1g c0q19 c0q20 c0q21 health cshn;
  run;

/* Subpopulation Analysis */
title2 'Multi-variate Freq: Overall Health by Asthma Status, Within Race';
proc surveyfreq;
  strata state;
  cluster idnumr;
  weight weight_1;
  table racec * c0q19 health / nocv;
  run;
```

The SAS System

Univariate Frequencies on Various Variables, All Children

The SURVEYFREQ Procedure

Data Summary

Number of Strata 51
Number of Clusters 192063
Number of Observations 192063
Sum of Weights 72786955

Is [S.C.] male or female?

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Std Err of Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - MALE</td>
<td>50564</td>
<td>51.1203</td>
<td>0.2734</td>
</tr>
<tr>
<td>2 - FEMALE</td>
<td>49719</td>
<td>48.8797</td>
<td>0.2734</td>
</tr>
<tr>
<td>Total</td>
<td>100270</td>
<td>100.000</td>
<td></td>
</tr>
</tbody>
</table>

Frequency Missing = 359

Race classification for all states [White, Black, Multiracial, Other]

<table>
<thead>
<tr>
<th>RACE</th>
<th>Frequency</th>
<th>Percent</th>
<th>Std Err of Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - WHITE ONLY</td>
<td>78403</td>
<td>74.8618</td>
<td>0.2788</td>
</tr>
<tr>
<td>2 - BLACK ONLY</td>
<td>10134</td>
<td>18.4491</td>
<td>0.2342</td>
</tr>
<tr>
<td>0 - MULTIPLE RACE</td>
<td>4407</td>
<td>3.7123</td>
<td>0.1040</td>
</tr>
<tr>
<td>4 - OTHER</td>
<td>4459</td>
<td>5.1674</td>
<td>0.1738</td>
</tr>
</tbody>
</table>

Has a doctor or health professional ever told you that [S.C.] has asthma?

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Std Err of Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - NO</td>
<td>89990</td>
<td>87.5454</td>
<td>0.1868</td>
</tr>
<tr>
<td>1 - YES</td>
<td>12202</td>
<td>12.4546</td>
<td>0.1868</td>
</tr>
<tr>
<td>Total</td>
<td>102192</td>
<td>100.000</td>
<td></td>
</tr>
</tbody>
</table>

Frequency Missing = 358
Subpopulation Analysis

Multi-variate Freqs: Overall Health by Asthma Status, Within Race

10:29 Saturday, December 2, 2006

The SURVEYFREQ Procedure

Data Summary

<table>
<thead>
<tr>
<th>Number of Strata</th>
<th>Number of Clusters</th>
<th>Number of Observations</th>
<th>Sum of Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>102353</td>
<td>102353</td>
<td>7276965</td>
</tr>
</tbody>
</table>

Table of S2019 by health

Controlling for BACEN=1, WHITE ONLY

<table>
<thead>
<tr>
<th>S2019</th>
<th>health</th>
<th>Frequency</th>
<th>Percent</th>
<th>Std Err of Percent</th>
<th>Row</th>
<th>Percent</th>
<th>Std Err of Row Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - NO</td>
<td>No</td>
<td>5632</td>
<td>8.5965</td>
<td>0.1863</td>
<td>9.7040</td>
<td>0.2093</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>82155</td>
<td>90.0103</td>
<td>0.2528</td>
<td>90.2880</td>
<td>0.2093</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>87787</td>
<td>98.6068</td>
<td>0.1849</td>
<td>100.000</td>
<td>0.2093</td>
<td></td>
</tr>
<tr>
<td>1 - YES</td>
<td>No</td>
<td>2001</td>
<td>3.0449</td>
<td>0.1066</td>
<td>26.7806</td>
<td>0.8201</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>6310</td>
<td>8.3482</td>
<td>0.1850</td>
<td>73.2594</td>
<td>0.8201</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>8311</td>
<td>11.3931</td>
<td>0.1849</td>
<td>100.000</td>
<td>0.8201</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>76237</td>
<td>100.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Logistic Regression:
SAS SURVEYLOGISTIC Syntax

- PROC SURVEYLOGISTIC;
  - STRATA STATE ;
  - CLUSTER IDNUMR ;
  - WEIGHT WEIGHT_I ;
  - CLASS IVAR1 (param=ref ref='VALUE');
  - MODEL DVAR = IVAR1 IVAR2 IVAR3;
  - FORMAT IVAR1 fmt1. IVAR2 fmt2.;
Logistic Regression: SAS SURVEYLOGISTIC Syntax

- PROC SURVEYLOGISTIC;
  - STRATA STATE ;
  - CLUSTER IDNUMR ;
  - WEIGHT WEIGHT_I ;
  - MODEL DVAR = IVAR1M IVAR2 IVAR3;
  - FORMAT IVAR1 fmt1. IVAR2 fmt2.;
data temp;
  set puf.schpuq_formatted;

*/ Create Analytical Variables */
/* qg56 - How would you describe condition of child's teeth */
/* qg01 - Male or Female: Male = 1 */
/* qg13 - Race classification: White, black, multiple, other */
/* qg14 - Is sample child of hispanic origin: Yes = 1 */

dvexg=-;
if qg56 in (1,2) then dvexg=7; /* Condition of teeth Excellent/Very Good */
if qg56 in (3,4,5) then dvexg=0; /* Condition of teeth Good, Fair, or Poor */

male=;
if qg01=1 then male=1; /* Male */
if qg01=2 then male=0; /* Female */

hisrace=;
if qg13=1 then hisrace=2; /* Non-hispanic white */
if qg13=2 then hisrace=9; /* Non-hispanic black */
if qg13 in (3,4) then hisrace=4; /* Non-hispanic other */
if qg14=1 then hisrace=1; /* Hispanic */

hispanic=; /* Dummy variable hispanic */
if hisrace=1 then hispanic=1;
if hisrace in (2,3,4) then hispanic=0;

nh_black=; /* Dummy variable non-hispanic black */
if hisrace=3 then nh_black=1;
if hisrace in (1,2,4) then nh_black=0;

nh_other=; /* Dummy variable non-hispanic other */
if hisrace=4 then nh_other=1;
if hisrace in (1,2,3) then nh_other=0;

/* Records Values to Missing */
if qg01 in (-M,6,7) then qg01 = .;

/* Create New Variable Labels */
label
  male  = 'Dummy variable for sex'
  hispanic = 'Dummy variable for hispanic'
  nh_black = 'Dummy variable for non-hispanic black'
  nh_other = 'Dummy variable for non-hispanic other'
  hisrace = 'Composite race/ethnicity variable'
  dvexg  = 'Child Dental Health is Excellent/Very Good'
;
run;
 PROC LOGISTIC DATA = texp;  
  strata state;  
  cluster idnum;  
  weight weight_1;  
  model dsexvg(descending) = agyr_child male hispanic nh_black nh_other;  
 RUN;  

 PROC LOGISTIC DATA = texp;  
  strata state;  
  cluster idnum;  
  weight weight_1;  
  model dsexvg(descending) = agyr_child female hispace(param=ref ref='2 - FEMALE') hispace(param=ref ref='NH White');  
  model dsexvg(descending) = agyr_child sex01 hispace;  
  format hispace hrace.;  
 RUN;  

/* Logistic Regression Using Dummy Variables */

/* Logistic Regression Using REFLEVEL Statement */

/* Creates Value Labels */

PROC FORMAT;  
  VALUE hrace  
1='Hispanic';  
2='NH White';  
3='NH Black';  
4='NH Other';  
RUN;
Oral Health Ex/VG by Demographics: binary dummy variables

The SURVEYLOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>4528247.48</td>
<td>6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>4628425.06</td>
<td>6</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>1977.0920</td>
<td>6</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1.6646</td>
<td>0.0332</td>
<td>2488.5080</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>AGEYR_CHILD</td>
<td>1</td>
<td>-0.0449</td>
<td>0.00204</td>
<td>2988.1002</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>male</td>
<td>1</td>
<td>-0.0988</td>
<td>0.0269</td>
<td>12.1980</td>
<td>0.0005</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td>-1.3627</td>
<td>0.0874</td>
<td>1380.1852</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>nh_black</td>
<td>1</td>
<td>0.7330</td>
<td>0.0881</td>
<td>864.0924</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>nh_other</td>
<td>1</td>
<td>-0.4400</td>
<td>0.0256</td>
<td>49.5933</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point Estimate</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGEYR_CHILD</td>
<td>0.955</td>
<td>0.951 - 0.961</td>
</tr>
<tr>
<td>male</td>
<td>0.510</td>
<td>0.504 - 0.516</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.255</td>
<td>0.250 - 0.260</td>
</tr>
<tr>
<td>nh_black</td>
<td>0.480</td>
<td>0.474 - 0.518</td>
</tr>
<tr>
<td>nh_other</td>
<td>0.644</td>
<td>0.637 - 0.753</td>
</tr>
</tbody>
</table>

Association of Predicted Probabilities and Observed Responses

<table>
<thead>
<tr>
<th></th>
<th>Percent Concordant</th>
<th>Somers' D</th>
<th>Percent Discordant</th>
<th>Gamma</th>
<th>Percent Tied</th>
<th>Tau-a</th>
<th>Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>61.7</td>
<td>0.245</td>
<td>36.6</td>
<td>0.261</td>
<td>2.0</td>
<td>0.097</td>
<td>1752880040</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The SAS System

Oral Health EX/VS by Demographics: Categorical Variables

The SURVEYLOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>4620247.40</td>
<td>6</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Score</td>
<td>4620445.06</td>
<td>6</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Wald</td>
<td>1677.8326</td>
<td>6</td>
<td>&lt; .0001</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Standard Estimate</th>
<th>Standard Error</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1.88101</td>
<td>0.32592</td>
<td>2488.6080</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>AGE69_CHILD</td>
<td>1</td>
<td>-0.04609</td>
<td>0.00284</td>
<td>288.1062</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>ST001 1 - MALE</td>
<td>1</td>
<td>-0.09380</td>
<td>0.02669</td>
<td>19.1100</td>
<td>0.0005</td>
</tr>
<tr>
<td>hisp = Hispanic</td>
<td>1</td>
<td>-1.5627</td>
<td>0.0874</td>
<td>1030.1962</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>hisp = NH Black</td>
<td>1</td>
<td>-0.7330</td>
<td>0.0884</td>
<td>564.9904</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>hisp = NH Other</td>
<td>1</td>
<td>0.6120</td>
<td>0.0626</td>
<td>10.6098</td>
<td>&lt; .0001</td>
</tr>
</tbody>
</table>

Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point Estimate</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE69_CHILD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST001 1 - MALE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hisp = Hispanic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hisp = NH Black</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hisp = NH Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Association of Predicted Probabilities and Observed Responses

<table>
<thead>
<tr>
<th>Percent Concordant</th>
<th>Somers' D</th>
<th>26.0</th>
<th>0.997</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Discordant</td>
<td>Gamma</td>
<td>0.028</td>
<td>0.097</td>
</tr>
<tr>
<td>Percent Tied</td>
<td>Tau-a</td>
<td>0.023</td>
<td>0.097</td>
</tr>
<tr>
<td>Pairs</td>
<td></td>
<td>1722880040</td>
<td></td>
</tr>
</tbody>
</table>
Frequency Distributions:
STATA SVY: TABULATE Syntax

- SVYSET [PWEIGHT=\text{WEIGHT}_I], STRATA(\text{STATE}) PSU(\text{IDNUMR})

- SVY:TABULATE V1 V2
  - Declare analysis variables
  - Declare the method used for variance estimation
  - Request specific table items such as standard errors, confidence limits, and row or column proportions
  - Request additional test statistics such as Chi-square or likelihood ratios
  - Define subgroup analyses
*** create analytical variables ***
*** Use Special Needs Screener to Classify a Child as Having a Special Need ***
generate canon = 0;
replace canon = 1 if s2q06==1 | s2q09==1 | s2q12==1 | s2q15==1 | s2q17==1;

" s2q06 - Prescription meds for a condition of 12 months or longer "
" s2q09 - Medical care, mental health, ed services for 12 months or longer "
" s2q12 - Limitation in abilities for a condition of 12 months or longer "
" s2q15 - Need special therapy for a condition for 12 months or longer "
" s2q17 - Emotional, behav problem for a condition for 12 months or longer "
generate health = .
replace health = 1 if s2q01==1 | s2q02==1;
replace health = 0 if s2q01==2 | s2q02==2 | s2q01==6;
generate ageg1 = .
replace ageg1 = 1 if aqyr_child <= 5;
replace ageg1 = 2 if aqyr_child >= 6 & aqyr_child <= 12;
replace ageg1 = 3 if aqyr_child <= 12 & aqyr_child <= 17;
* s2q09 - Male or Female
* s2q12 - Ever told by doctor had asthma
* s2q13 - Ever told by doctor had hearing or vision problems
* s2q14 - Ever told by doctor had ADHD
* racer - Race classification
* aqyr08r - How many times gone to doctor for sick care

*** Recode values to missing ***
recode s1q01 6 7 8 9 0 . l . m . n = .
recode s2q01 6 7 8 9 0 . l . m . n = .
recode s2q02 6 7 8 9 0 . l . m . n = .
recode s2q03 6 7 8 9 0 . l . m . n = .
recode s2q05 6 7 8 9 0 . l . m . n = .
recode racer 6 7 8 9 0 . l . m . n = .
recode s4q01 98 97 . l . m . n = .

*** Create new variable labels ***
label variable canon "Child has Special Health Care Needs"
label variable health 'Child Overall Health is Excellent/Very Good'
label variable ageg1 "Age of Child Grouping !"

*** Create variable value formats ***
label define ynesa
  0 "No"
  1 "Yes"
label define ageg1
  1 "0-5 Years"
  2 "6-11 Years"
  3 "12-17 Years"
label values health ynesa;
label values ynesa;
label values ageg1 ageg1;

svyset founum [pweight=weight_1], strata(state) vce(linearized)

pweight: weight_1
vce: linearized
Strata 1: state
SU 1: founum
PSU 1: <zero>
svyset idnum | pweight

pweight: weight_1
VCE: lineariz
strata: state
SU: <0num
FPC 1: <zero>

svytabulate - One-way tables for survey data

Model  |  f  |  SE  |  Tado  |  Respoint |
-------|----|-----|-------|-----------|
Cell proportions
Weighted cell counts
Standard errors
Confidence intervals
Display the DEFT design effects
Display the DEFT design effects
Repeat design effects assuming SRS within subpopulation
Cell observations

Command
. svy, vce(linearized): tabulate race, se of percent
(running tabulate on estimation sample)

<table>
<thead>
<tr>
<th>Race classification for all states (White, Black, Multiracial, Other)</th>
<th>percentages</th>
<th>se</th>
<th>lb</th>
<th>ub</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - White</td>
<td>74.65</td>
<td>.2788</td>
<td>74.1</td>
<td>75.19</td>
</tr>
<tr>
<td>2 - Black</td>
<td>15.45</td>
<td>.2342</td>
<td>15</td>
<td>16.92</td>
</tr>
<tr>
<td>3 - Mult</td>
<td>3.712</td>
<td>.204</td>
<td>3.514</td>
<td>3.922</td>
</tr>
<tr>
<td>4 - Other</td>
<td>5.187</td>
<td>.1796</td>
<td>4.847</td>
<td>5.511</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key:
- percentages = cell percentages
- se = linearized standard errors of cell percentages
- lb = lower 95% confidence bounds for cell percentages
- ub = upper 95% confidence bounds for cell percentages
Number of strata = 9
Number of PSUs = 614

<table>
<thead>
<tr>
<th>Race Classification for All States (White, Black, Asian, Hispanic, Other)</th>
<th>percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - White</td>
<td>74.65%</td>
</tr>
<tr>
<td>2 - Black</td>
<td>10.45%</td>
</tr>
<tr>
<td>3 - Asian</td>
<td>3.71%</td>
</tr>
<tr>
<td>4 - Other</td>
<td>11.18%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>

Variables:
- s8001: During the past 12 months, would you say that...
- s9000: Would you say that...
- s9100: Would you say that...
- s9110: Would you say that...
- s9119: Would you say that...
- s9120: Would you say that...
- s915: During the past 12 months, do you have any...
- s915b: During the past 12 months, do you have any...
- s915c: (Do you have any...)
- s10001: People in this neighborhood, do you watch out for...
- s10002: We watch out for...
- s10003: There are people...
- s10004: There are people...
- s10005: If my child ever is...
- s10006: How often do you...
- s10007: How often do you...
- s10008: How often do you...
- s10009: Is (S.C.) in the neighborhood...?
Logistic Regression: STATA SVY:LOGIT

- SVYSET [PWEIGHT=\textit{WEIGHT}_I], STRATA(\textit{STATE}) PSU(\textit{IDNUMR})

- SVY: LOGIT \textit{DVAR1 IVAR1 IVAR2 IVAR3}, OR
  - Declare dependent analysis variable
  - Define subgroup analyses
  - Declare the method used for variance estimation
  - Declare reporting specifications such as the confidence interval levels and request odds ratios or dependent variable coefficients
  - Define the maximum number or iterations to run
*** Create analytical variables ***
```
** s2q64 - How would you describe condition of child's teeth **
** s1q01 - Male or female: Male = 1 **
** racec - Race classification: White, black, multiple, other **
** s1lq01 - Is sample child of Hispanic origin: Yes = 1 **

generate duexvg = .;
replace duexvg = 1 if s2q64==1 | s2q64==2;
replace duexvg = 0 if s2q64==3 | s2q64==4 | s2q64==5;

generate male = .;
replace male = 1 if s1q01==1;
replace male = 0 if s1q01==2;

generate hisracec = .;
replace hisracec = 2 if racec == 1;
replace hisracec = 3 if racec == 2;
replace hisracec = 4 if racec == 3 | racec == 4;
replace hisracec = 5 if s1lq01 == 1;

generate hispanic = .;
replace hispanic = 1 if hisracec == 1;
replace hispanic = 0 if hisracec == 2 | hisracec == 5 | hisracec == 4;

generate nh_black = .;
replace nh_black = 1 if hisracec == 3;
replace nh_black = 0 if hisracec == 1 | hisracec == 4 | hisracec == 5;

generate nh_other = .;
replace nh_other = 1 if hisracec == 4;
replace nh_other = 0 if hisracec == 1 | hisracec == 2 | hisracec == 3;

*** Recode Values to Missing ***
recode s1lq01 6/7 . 8/9 .
```

*** Create new variable labels ***
```
label variable male "Dummy variable for males";
label variable hispanic "Dummy variable for hispanics";
label variable nh_black "Dummy variable for non-Hispanic blacks";
label variable nh_other "Dummy variable for non-Hispanic others";
label variable hisracec "Composite race ethnicity variable"
label variable duexvg "Child Dental Health is Excellent/Very Good"
```

*** Create variable value for racec ***
```
label define racec 1 'Hispanic' 2 'NH White' 3 'NH Black' 4 'NH Other'
```
```
label values hisracec racec;
```
```
.logistic dvexvg ageyr_child male Hispanic nh_black nh_other, coef
(running logistic on estimation sample)

Survey: Logistic regression
Number of strata = 51
Number of obs = 94292
Population size = 67310349
Design df = 94241
F(  5, 94237) = 333.37
Prob > F = 0.0000

|        | Coef. | Std. Err. | t     | P>|t|  | [95% Conf. Interval] |
|--------|-------|-----------|-------|-----|----------------------|
| dvexvg | -0.0408572 | 0.0016427 | -24.97 | 0.000 | -0.0502599 | -0.029327 |
| ageyr_child | -0.0937826 | 0.0045518 | -20.38 | 0.000 | -0.1379425 | -0.0452392 |
| male         | -1.1602556 | 0.0178216 | -64.97 | 0.000 | -1.2974947 | -0.9220168 |
| Hispanic      | -0.7328664 | 0.0238683 | -30.11 | 0.000 | -0.8168534 | -0.6488793 |
| nh_black      | -0.0400972 | 0.0082817 | -4.84 | 0.000 | -0.1324789 | -0.0477073 |
| nh_other      | 1.654497  | 0.0131654 | 49.89 | 0.000 | 1.5994933 | 1.719501 |
```

Variables:
- idnum: Unique ID number
- state: State of residence
- meta_stat: Metropolitan Status
- ageyr_child: Desired Age/yr
- total_area: Total area
- agepos4: Age position of the
- is_s: Is S.C. male or?
- relation: Desired Relation
- total_area: Total number of a
- educ_num: What is the higher
- plang: What is the prime
- high_imp: Flag indicating sex
- s2p3: How much does
- bmi: BMI
- is_s: Does S.C. cause
- is_n: Is [i/s] need
- is_a: Is this a condition
- is_s: Does S.C. need
- is_n: Is [i/s] need
- is_a: Is this a condition
- is_s: Does S.C. need
- is_n: Is [i/s] need
- is_a: Is this a condition
- is_s: Does S.C. need
- is_n: Is [i/s] need
- is_a: Is this a condition
svy, vce(linearized): logit dvexy ageyr_child male hispanic rh_black rh_other, or
(running logit on estimation sample)

Survey: Logistic regression

| Variable | Odds Ratio | Std. Err. | t | P>|t| | [95% Conf. Interval] |
|----------|------------|-----------|---|------|----------------------|
| ageyr_child | .954134 | .0025269 | -16.97 | 0.000 | .949627 | .958668 |
| Hispanic | .2339646 | .0095635 | -26.47 | 0.000 | .223793 | .244133 |
| rh_black | .4904972 | .0184393 | -21.11 | 0.000 | .4649622 | .5179995 |
| rh_other | .9240515 | .0604409 | -15.06 | 0.000 | .8679056 | .9805469 |
For More Information…

- Julian Luke or Stephen Blumberg
  Centers for Disease Control and Prevention
  National Center for Health Statistics
  3311 Toledo Road
  Hyattsville, Maryland 20782 USA

- jluke@cdc.gov or sblumberg@cdc.gov