MULTILOG Example #5

SUDAAN Statements and Results Illustrated

- PREDMARG
- ADJRR option
- GENLOGIT option
- PRED_EFF
- SUBPOPX

Input Data Set(s): NH3MI1.SAS7bdat - NH3MI5.SAS7bdat

Example

Using data from the NHANES III Multiply Imputed (MI) dataset, model the effects of age, weight, gender, and current smoking status among adults age 20 yrs and older on the self-rating of health status.

In addition to featuring analysis of multiply imputed data, this example also highlights the estimation of model-adjusted risks, risk ratios, and risk differences for a multinomial outcome via predicted marginal proportions (ADJRR option on PREDMARG statement). Confidence intervals for model-adjusted risks (marginals) are new in Release 11.0.

Solution

This example illustrates the MULTILOG SUDAAN procedure with the following three multiply imputed variables.

HAB1MI = Self-rating of health status. This is a five-level categorical variable coded 1=Excellent, 2=Very good, 3=Good, 4=Fair, and 5=Poor.

HAM6MI = Weight in pounds (continuous).

HAR3RMI = Current smoking status (1=yes, 2=no).

The model also uses AGE (continuous, in years) and HSSEX (1=Male, 2=Female) as independent variables. Only one version of these variables is available. These two variables, along with all design statement variables, are repeated in *each* of the five MI datasets.

Five different versions of these imputation-revised variables were constructed and saved on the *Nh3mi1*,…,*Nh3mi5* SAS datasets, respectively. In the DATA option on the PROC MULTILOG statement, the user only needs to specify the *first* dataset in the series of five. The MI_COUNT=5 option tells SUDAAN that there are five MI datasets available for analysis, numbered sequentially.

The NEST statement indicates that SDPSTRA6 is the stratification variable and SDPPSU6 is the primary sampling unit. The WEIGHT statement indicates that WTPFQX6 contains the analysis weight for each record on the file.

The SUBPOPX statement restricts the analysis to adults aged 20 years or older.

In MULTILOG, the CLASS statement contains the dependent variable *and* all covariates that are to be modeled as categorical (HAB1MI is the dependent; HSSEX and HAR3RMI are the categorical independent).

The GENLOGIT (generalized logit) link specifies the generalized logit model. The default Wald-*F* test is used for all tests of hypotheses.

The PREDMARG statement requests the predicted marginal proportion (*model-adjusted risk*) for each level of HSSEX. The ADJRR option on the PREDMARG statement computes the *model-adjusted risk ratio* for males vs. females. Finally, the PRED_EFF statement requests the *difference* in predicted marginal proportions (*risk differences*) for Males vs. Females.

We include multiple PRINT statements, all of which are optional. Multiple PRINT statements allow the user to set up different default print environments (SETENV statements) for different PRINT groups. The PRINT statements are used in this example to request the PRINT groups of interest; to calculate individual statistics of interest, and in some cases, change default labels for those statistics; and to specify a variety of formats for those printed statistics. Without the PRINT statement, default statistics are produced from each PRINT group, with default formats.

The SETENV statements are optional. They set up default formats for printed statistics and further manipulate the printout to the needs of the user.

The RFORMAT and RLABEL statements associate SAS formats and variable labels, respectively, with the variables used in the procedure. Without the RLABEL statement, SAS variable labels would be produced if already defined.

This example was run in SAS-Callable SUDAAN, and the SAS program and *.LST files are provided.

Exhibit 1. SAS-Callable SUDAAN Code

```
libname in "c:\903winbetatest\nhanes3";
options nocenter linesize=85 pagesize=68;

proc format;
  value hab 1="1=Excellent"
            2="2=Very Good"
            3="3=Good"
            4="4=Fair"
            5="5=Poor";
  value sex 1="1=Male"
            2="2=Female";
  value smoke 1="1=Yes"
              2="2=No";

data mi1; set in.nh3mi1;
   AGE = HSAGEIR;
   IF HSAGEU = 1 THEN AGE = AGE / 12;
   keep sdpstra6 sdppsu6 wtpfqx6 age hssex hab1mi ham6mi har3rmi;
proc sort data=mi1; by SDPSTRA6 SDPPSU6;

data mi2; set in.nh3mi2;
   AGE = HSAGEIR;
   IF HSAGEU = 1 THEN AGE = AGE / 12;
   keep sdpstra6 sdppsu6 wtpfqx6 age hssex hab1mi ham6mi har3rmi;
proc sort data=mi2; by SDPSTRA6 SDPPSU6;

data mi3; set in.nh3mi3;
    AGE = HSAGEIR;
    IF HSAGEU = 1 THEN AGE = AGE / 12;
    keep sdpstra6 sdppsu6 wtpfqx6 age hssex hab1mi ham6mi har3rmi;
proc sort data=mi3; by SDPSTRA6 SDPPSU6;

data mi4; set in.nh3mi4;
     AGE = HSAGEIR;
     IF HSAGEU = 1 THEN AGE = AGE / 12;
     keep sdpstra6 sdppsu6 wtpfqx6 age hssex hab1mi ham6mi har3rmi;
proc sort data=mi4; by SDPSTRA6 SDPPSU6;

data mi5; set in.nh3mi5;
      AGE = HSAGEIR;
      IF HSAGEU = 1 THEN AGE = AGE / 12;
      keep sdpstra6 sdppsu6 wtpfqx6 age hssex hab1mi ham6mi har3rmi;
proc sort data=mi5; by SDPSTRA6 SDPPSU6;
```

Exhibit 1. SAS-Callable SUDAAN Code-cont.

```
PROC MULTILOG DATA=mi1 FILETYPE=SAS mi_count=5 DESIGN=WR;
  NEST SDPSTRA6 SDPPSU6 / MISSUNIT;
  WEIGHT WTPFQX6;

  SUBPOPX age>=20 / name="Age 20+";
  CLASS HAB1MI HSSEX HAR3RMI;
  MODEL HAB1MI = AGE HAM6MI HSSEX HAR3RMI / genlogit;
  PREDMARG HSSEX / adjrr;
  PRED_EFF HSSEX=(1 -1) / name="Males-Females";

  setenv labwidth=25 colspce=1 topmgn=0 colwidth=7 decwidth=4;
  print / betas=default risk=default tests=default t_betafmt=f6.2 waldffmt=f6.2
          waldpfmt=f7.4 dffmt=f7.0 orfmt=f10.3 loworfmt=f9.3 uporfmt=f9.3
          style=nchs;

  SETENV LABWIDTH=25 COLWIDTH=7 DECWIDTH=4 COLSPCE=3 TOPMGN=0;
  PRINT / PRED_MRG=default PRMGCONS=default lowpmfmt=f6.4 uppmfmt=f6.4
          predmrgfmt=f9.4 t_prdmrgfmt=f8.2 prmgconfmt=f8.4 t_pmconfmt=f7.2;

  SETENV LABWIDTH=30 COLWIDTH=5 DECWIDTH=3 COLSPCE=2 TOPMGN=0;
  PRINT PRED_RR="Risk Ratio" / PREDRISK=default;

  RLABEL HAB1MI="Self-Rating Health Status";
  RLABEL HAM6MI="Weight (lbs)";
  RLABEL HAR3RMI="Current Smoker?";

  RFORMAT hab1mi hab.;
  RFORMAT hssex sex.;
  RFORMAT har3rmi smoke.;
  RTITLE "Modelling Self-Rating of Health Status (MI Data)";
  RFOOTNOTE "NHANES III, Multiply Imputed Dataset:  Respondents Age 20+";
```

Exhibit 2. First Page of SUDAAN Output

```
                              S U D A A N
           Software for the Statistical Analysis of Correlated Data
          Copyright      Research Triangle Institute    November 2011
                               Release 11.0.0


DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method,
Assuming a With Replacement (WR) Design
    Sample Weight: WTPFQX6
    Stratification Variables(s): SDPSTRA6
    Primary Sampling Unit: SDPPSU6


Processing data for set 1 of imputed variables:

Processing data for set 2 of imputed variables:

Processing data for set 3 of imputed variables:

Processing data for set 4 of imputed variables:

Processing data for set 5 of imputed variables:


Processing data for set 1 of imputed variables:

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     49


Maximum number of estimable parameters for the model is 20

File MI1 contains   98 Clusters
  98 clusters were used to fit the model
Maximum cluster size is 297 records
Minimum cluster size is  75 records


Sample and Population Counts for Response Variable HAB1MI
Based on observations used in the analysis
  1=Excellent:  Sample Count     2824    Population Count  36794794
  2=Very Good:  Sample Count     4387    Population Count  54082074
  3=Good     :  Sample Count     6743    Population Count  57598850
  4=Fair     :  Sample Count     3833    Population Count  22784573
  5=Poor     :  Sample Count     1038    Population Count   5920379
```

Exhibit 2. First Page of SUDAAN Output-cont.

```
Processing data for set 2 of imputed variables:

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     49


Maximum number of estimable parameters for the model is 20

File MI2 contains   98 Clusters
  98 clusters were used to fit the model
Maximum cluster size is 297 records
Minimum cluster size is  75 records


Sample and Population Counts for Response Variable HAB1MI
Based on observations used in the analysis
  1=Excellent:  Sample Count     2823    Population Count  36789793
  2=Very Good:  Sample Count     4389    Population Count  54114095
  3=Good     :  Sample Count     6743    Population Count  57609798
  4=Fair     :  Sample Count     3833    Population Count  22757928
  5=Poor     :  Sample Count     1037    Population Count   5909056


Processing data for set 3 of imputed variables:

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     49


Maximum number of estimable parameters for the model is 20

File MI3 contains   98 Clusters
  98 clusters were used to fit the model
Maximum cluster size is 297 records
Minimum cluster size is  75 records


Sample and Population Counts for Response Variable HAB1MI
Based on observations used in the analysis
  1=Excellent:  Sample Count     2823    Population Count  36789793
  2=Very Good:  Sample Count     4390    Population Count  54095547
  3=Good     :  Sample Count     6740    Population Count  57614324
  4=Fair     :  Sample Count     3835    Population Count  22771679
  5=Poor     :  Sample Count     1037    Population Count   5909327
```

Exhibit 2. First Page of SUDAAN Output-cont.

```
Processing data for set 4 of imputed variables:

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     49


Maximum number of estimable parameters for the model is 20

File MI4 contains   98 Clusters
  98 clusters were used to fit the model
Maximum cluster size is 297 records
Minimum cluster size is  75 records


Sample and Population Counts for Response Variable HAB1MI
Based on observations used in the analysis
  1=Excellent:  Sample Count     2824    Population Count  36791051
  2=Very Good:  Sample Count     4388    Population Count  54095607
  3=Good     :  Sample Count     6738    Population Count  57585236
  4=Fair     :  Sample Count     3837    Population Count  22788126
  5=Poor     :  Sample Count     1038    Population Count   5920649


Processing data for set 5 of imputed variables:

Independence parameters have converged in 6 iterations

Number of observations read       :  33994    Weighted count:251097002
Observations in subpopulation     :  18825    Weighted count:177180670
Observations used in the analysis :  18825    Weighted count:177180670
Denominator degrees of freedom    :     49


Maximum number of estimable parameters for the model is 20

File MI5 contains   98 Clusters
  98 clusters were used to fit the model
Maximum cluster size is 297 records
Minimum cluster size is  75 records


Sample and Population Counts for Response Variable HAB1MI
Based on observations used in the analysis
  1=Excellent:  Sample Count     2824    Population Count  36808280
  2=Very Good:  Sample Count     4387    Population Count  54082697
  3=Good     :  Sample Count     6741    Population Count  57602295
  4=Fair     :  Sample Count     3836    Population Count  22778341
  5=Poor     :  Sample Count     1037    Population Count   5909056

Overall degrees of freedom (Rubin): 46.36

-2 * Normalized Log-Likelihood with Intercepts Only : 53880.98
-2 * Normalized Log-Likelihood Full Model           : 51917.42
Approximate Chi-Square (-2 * Log-L Ratio)           :  1963.56
Degrees of Freedom                                  :       16

Note: The approximate Chi-Square is not adjusted for clustering.
      Refer to hypothesis test table for adjusted test.
```

The sample design information is followed by summary information on each of the imputed datasets: convergence of parameters; observations read in subpopulation and used in analysis; number of clusters; minimum and maximum cluster size; and a distribution of the response variable based on observations used in the analysis (***Exhibit 2***). In each dataset, there were 33,994 observations read, with 18,825 used

in the analysis. The distribution of the response variable changed slightly from dataset to dataset, since HAB1MI is one of the multiply imputed variables in the analysis.

The following results represent the summary over all imputations (the option BY_MI on the PRINT statement produces results separately for each imputation). The frequency distributions for variables on the CLASS statement are presented first (***Exhibit 3*** to ***Exhibit 5***).

Exhibit 3. CLASS Variable Frequencies (HAB1MI)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Self-Rating Health Status.
-------------------------------------------
Self-Rating
  Health
  Status            Frequency          Value
-------------------------------------------
Ordered
  Position:
  1                  2823.600    1=Excellent
Ordered
  Position:
  2                  4388.200    2=Very Good
Ordered
  Position:
  3                  6741.000         3=Good
Ordered
  Position:
  4                  3834.800         4=Fair
Ordered
  Position:
  5                  1037.400         5=Poor
-------------------------------------------
```

Exhibit 4. CLASS Variable Frequencies (Sex)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Sex.
-----------------------------------
Sex             Frequency       Value
-----------------------------------
Ordered
  Position:
  1                  8816      1=Male
Ordered
  Position:
  2                 10009    2=Female
-----------------------------------
```

Exhibit 5. CLASS Variable Frequencies (Current Smoker?)

```
Frequencies and Values for CLASS Variables
Results for Summary Over All Imputations

by: Current Smoker?.
--------------------------------------
Current
  Smoker?           Frequency    Value
--------------------------------------
Ordered
  Position:
  1                  4804.200    1=Yes
Ordered
  Position:
  2                 14020.800     2=No
--------------------------------------
```

Exhibit 6. Estimated Regression Coefficients

Variance Estimation Method: Taylor Series (WR) Using Multiply Imputed Data
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable HAB1MI: Self-Rating Health Status
For Subpopulation: Age 20+

Modelling Self-Rating of Health Status (MI Data)

Results for Summary Over All Imputations

HAB1MI (log-odds) Independent Variables and Effects	Beta Coeff.	SE Beta	Lower 95% Limit Beta	Upper 95% Limit Beta	T-Test B=0	P-value T-Test B=0
1=Excellent vs 5=Poor						
Intercept	7.1332	0.4295	6.2691	7.9974	16.61	0.0000
Sex						
1=Male	0.6362	0.1125	0.4099	0.8625	5.65	0.0000
2=Female	0.0000	0.0000	.	.	.	.
Current Smoker?						
1=Yes	-1.2645	0.1637	-1.5938	-0.9351	-7.72	0.0000
2=No	0.0000	0.0000	.	.	.	.
AGE	-0.0701	0.0036	-0.0773	-0.0629	-19.60	0.0000
Weight (lbs)	-0.0104	0.0018	-0.0140	-0.0069	-5.90	0.0000
2=Very Good vs 5=Poor						
Intercept	6.4642	0.4335	5.5919	7.3365	14.91	0.0000
Sex						
1=Male	0.3354	0.1046	0.1251	0.5458	3.21	0.0024
2=Female	0.0000	0.0000	.	.	.	.
Current Smoker?						
1=Yes	-0.8281	0.1563	-1.1424	-0.5137	-5.30	0.0000
2=No	0.0000	0.0000	.	.	.	.
AGE	-0.0658	0.0036	-0.0731	-0.0585	-18.05	0.0000
Weight (lbs)	-0.0047	0.0019	-0.0084	-0.0009	-2.52	0.0152
3=Good vs 5=Poor						
Intercept	5.3203	0.3660	4.5839	6.0566	14.54	0.0000
Sex						
1=Male	0.1541	0.1138	-0.0748	0.3829	1.35	0.1822
2=Female	0.0000	0.0000	.	.	.	.
Current Smoker?						
1=Yes	-0.4394	0.1574	-0.7561	-0.1227	-2.79	0.0076
2=No	0.0000	0.0000	.	.	.	.
AGE	-0.0536	0.0031	-0.0599	-0.0474	-17.27	0.0000
Weight (lbs)	-0.0008	0.0017	-0.0042	0.0027	-0.45	0.6577
4=Fair vs 5=Poor						
Intercept	2.8866	0.4232	2.0352	3.7380	6.82	0.0000
Sex						
1=Male	-0.1220	0.1292	-0.3819	0.1379	-0.94	0.3497
2=Female	0.0000	0.0000	.	.	.	.
Current Smoker?						
1=Yes	-0.1732	0.1515	-0.4780	0.1316	-1.14	0.2588
2=No	0.0000	0.0000	.	.	.	.
AGE	-0.0312	0.0034	-0.0380	-0.0244	-9.25	0.0000
Weight (lbs)	0.0020	0.0017	-0.0015	0.0054	1.15	0.2563

NHANES III, Multiply Imputed Dataset: Respondents Age 20+

The generalized logit model fits a separate logit equation for *each* of the response categories vs. the last (***Exhibit 6***). Males have a significantly increased chance of being in the Excellent self-reported health status category compared to women. Being a non-smoker, as well as having lowered weight and age, are also significantly associated with self-reporting Excellent or Very Good health status.

Exhibit 7. ANOVA Table

```
Variance Estimation Method: Taylor Series (WR) Using Multiply Imputed Data
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable HAB1MI: Self-Rating Health Status
For Subpopulation: Age 20+

Modelling Self-Rating of Health Status (MI Data)

Results for Summary Over All Imputations

-----------------------------------------------------------------
Contrast                        Degrees
                                of                      P-value
                                Freedom       Wald F    Wald F
-----------------------------------------------------------------
OVERALL MODEL                        20       146.23     0.0000
MODEL MINUS INTERCEPT                16        53.21     0.0000
INTERCEPT                             .           .          .
HSSEX                                 4        32.46     0.0000
HAR3RMI                               4        37.85     0.0000
AGE                                   4       122.46     0.0000
HAM6MI                                4        33.23     0.0000
-----------------------------------------------------------------
NHANES III, Multiply Imputed Dataset:  Respondents Age 20+
```

Exhibit 7 indicates that all four independent variables (sex, current smoker, age, and weight) are significantly associated with HAB1MI, self-rating of general health status among people 20 years of age and older. Note that each variable effect has four df, since there are four logit equations.

Exhibit 8. Default Odds Ratios

```
Variance Estimation Method: Taylor Series (WR) Using Multiply Imputed Data
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable HAB1MI: Self-Rating Health Status
For Subpopulation: Age 20+

Modelling Self-Rating of Health Status (MI Data)

Results for Summary Over All Imputations

------------------------------------------------------------
HAB1MI (log-odds)
   Independent Variables                 Lower 95%   Upper 95%
     and Effects            Odds Ratio   Limit OR    Limit OR
------------------------------------------------------------
1=Excellent vs 5=Poor
   Intercept                1252.933     527.997    2973.201
   Sex
     1=Male                    1.889       1.507       2.369
     2=Female                  1.000         .           .
   Current Smoker?
     1=Yes                     0.282       0.203       0.393
     2=No                      1.000         .           .
   AGE                         0.932       0.926       0.939
   Weight (lbs)                0.990       0.986       0.993
2=Very Good vs 5=Poor
   Intercept                 641.755     268.241    1535.372
   Sex
     1=Male                    1.399       1.133       1.726
     2=Female                  1.000         .           .
   Current Smoker?
     1=Yes                     0.437       0.319       0.598
     2=No                      1.000         .           .
   AGE                         0.936       0.929       0.943
   Weight (lbs)                0.995       0.992       0.999
3=Good vs 5=Poor
   Intercept                 204.441      97.897     426.940
   Sex
     1=Male                    1.167       0.928       1.467
     2=Female                  1.000         .           .
   Current Smoker?
     1=Yes                     0.644       0.469       0.885
     2=No                      1.000         .           .
   AGE                         0.948       0.942       0.954
   Weight (lbs)                0.999       0.996       1.003
4=Fair vs 5=Poor
   Intercept                  17.932       7.654      42.013
   Sex
     1=Male                    0.885       0.683       1.148
     2=Female                  1.000         .           .
   Current Smoker?
     1=Yes                     0.841       0.620       1.141
     2=No                      1.000         .           .
   AGE                         0.969       0.963       0.976
   Weight (lbs)                1.002       0.999       1.005
------------------------------------------------------------
NHANES III, Multiply Imputed Dataset:  Respondents Age 20+
```

Some interpretations from the table of default odds ratios in ***Exhibit 8***: the odds of self-reporting Excellent or Very Good health are increased by 89% (OR=1.889) and 40% (OR=1.399), respectively, in males compared to females. Being a current smoker reduces the odds of self-reporting Excellent health by more than 70% (OR=0.282).

Exhibit 9. Predicted Marginal Proportions (Model-Adjusted Risks)

Variance Estimation Method: Taylor Series (WR) Using Multiply Imputed Data
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable HAB1MI: Self-Rating Health Status
For Subpopulation: Age 20+

Modelling Self-Rating of Health Status (MI Data)

Results for Summary Over All Imputations

Self-Rating Health Status

Predicted Marginal #1	Predicted Marginal	SE	Lower 95% Limit	Upper 95% Limit	T:Marg=0	P-value
1=Excellent						
Sex						
1=Male	0.2477	0.0089	0.2302	0.2660	27.78	0.0000
2=Female	0.1760	0.0073	0.1618	0.1912	24.11	0.0000
2=Very Good						
Sex						
1=Male	0.3120	0.0077	0.2966	0.3278	40.29	0.0000
2=Female	0.2953	0.0093	0.2769	0.3144	31.67	0.0000
3=Good						
Sex						
1=Male	0.3055	0.0073	0.2910	0.3205	41.70	0.0000
2=Female	0.3406	0.0093	0.3221	0.3596	36.54	0.0000
4=Fair						
Sex						
1=Male	0.1054	0.0055	0.0949	0.1169	19.33	0.0000
2=Female	0.1514	0.0079	0.1363	0.1679	19.28	0.0000
5=Poor						
Sex						
1=Male	0.0294	0.0023	0.0252	0.0343	13.07	0.0000
2=Female	0.0366	0.0025	0.0320	0.0419	14.85	0.0000

NHANES III, Multiply Imputed Dataset: Respondents Age 20+

Exhibit 9 presents the predicted marginal proportions (model-adjusted risks) and their 95% confidence limits for each category of health status, separately for each sex. It becomes clear that men are more likely to report being in Excellent health than are women (25% vs. 18%, respectively, with confidence limits that do not overlap). Women are more likely to report being in the less desirable categories of Good, Fair, or Poor health compared to men. This is after adjusting for age, weight, and current smoking status.

Exhibit 10. Model-Adjusted Risk Differences

Variance Estimation Method: Taylor Series (WR) Using Multiply Imputed Data
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable HAB1MI: Self-Rating Health Status
For Subpopulation: Age 20+

Modelling Self-Rating of Health Status (MI Data)

Results for Summary Over All Imputations

Self-Rating Health Status Contrasted Predicted Marginal #1	PREDMARG Contrast	SE	T-Stat	P-value
1=Excellent				
Males-Females	0.0716	0.0080	8.94	0.0000
2=Very Good				
Males-Females	0.0166	0.0103	1.61	0.1140
3=Good				
Males-Females	-0.0351	0.0091	-3.84	0.0004
4=Fair				
Males-Females	-0.0460	0.0070	-6.54	0.0000
5=Poor				
Males-Females	-0.0072	0.0033	-2.18	0.0344

NHANES III, Multiply Imputed Dataset: Respondents Age 20+

Exhibit 10 shows that the difference in male vs. female risk is significant for all but one category of response. Men are significantly more prevalent than women in the Excellent health category, while women are significantly more prevalent than men in the Good, Fair, and Poor health categories.

Exhibit 11. Model-Adjusted Risk Ratios

```
Variance Estimation Method: Taylor Series (WR) Using Multiply Imputed Data
SE Method: Robust (Binder, 1983)
Working Correlations: Independent
Link Function: Generalized Logit
Response variable HAB1MI: Self-Rating Health Status
For Subpopulation: Age 20+

Modelling Self-Rating of Health Status (MI Data)

Results for Summary Over All Imputations

-----------------------------------------------------------------
Self-Rating Health Status
   Predicted Marginal Risk                        Lower   Upper
     Ratio #1                    Risk              95%     95%
                                 Ratio      SE   Limit   Limit
-----------------------------------------------------------------
1=Excellent
   Sex
     1=Male vs. 2=Female         1.407   0.053   1.303   1.519
2=Very Good
   Sex
     1=Male vs. 2=Female         1.056   0.036   0.986   1.132
3=Good
   Sex
     1=Male vs. 2=Female         0.897   0.025   0.848   0.949
4=Fair
   Sex
     1=Male vs. 2=Female         0.696   0.036   0.627   0.773
5=Poor
   Sex
     1=Male vs. 2=Female         0.804   0.081   0.657   0.984
-----------------------------------------------------------------
NHANES III, Multiply Imputed Dataset:  Respondents Age 20+
```

The ratio of male vs. female risk is presented in ***Exhibit 11***, above. The risk of self-reported Excellent health is increased by 41% (RR=1.407) for men vs. women. Conversely, the risk of being in Good, Fair, or Poor health is reduced by 10%-30% in men vs. women (RR=0.897, 0.696, and 0.804 for Good, Fair, and Poor health, respectively). All but one of the 95% confidence limits do not contain the null value of 1.0.