The data are from a longitudinal clinical trial of contracepting women. In this trial women received an injection of either 100 mg or 150 mg of depot-medroxyprogesterone acetate (DMPA) on the day of randomization and three additional injections at 90-day intervals. There was a final follow-up visit 90 days after the fourth injection, i.e., one year after the first injection.

Throughout the study each woman completed a menstrual diary that recorded any vaginal bleeding pattern disturbances. The diary data were used to determine whether a women experienced amenorrhea, the absence of menstrual bleeding for a specified number of days.

A total of 1151 women completed the menstrual diaries and the diary data were used to generate a binary sequence for each woman according to whether or not she had experienced amenorrhea in the four successive three month intervals.

Reference: Machin D, Farley T, Busca B, Campbell M and d'Arcangues C. (1988). *Assessing changes in vaginal bleeding patterns in contracepting women*. **Contraception**, 38, 165-179.

**proc** **print** data = BIOS755.Amenorrhea (obs=**10**);

**run**;

|  |
| --- |
| The SAS System |

| **Obs** | **ID** | **TRT** | **TIME** | **Y** | **Ctime** | **prevy** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | 1 | 0 | 1 | 0 | 1 | . |
| **2** | 1 | 0 | 2 | . | 2 | 0 |
| **3** | 1 | 0 | 3 | . | 3 | . |
| **4** | 1 | 0 | 4 | . | 4 | . |
| **5** | 2 | 0 | 1 | 0 | 1 | . |
| **6** | 2 | 0 | 2 | . | 2 | 0 |
| **7** | 2 | 0 | 3 | . | 3 | . |
| **8** | 2 | 0 | 4 | . | 4 | . |
| **9** | 3 | 0 | 1 | 0 | 1 | . |
| **10** | 3 | 0 | 2 | . | 2 | 0 |

**First, we are going to look at two different ways of imputing and the results they give. The first way uses a monotone statement on the long version of the dataset.**

**proc** **mi** data=BIOS755.Amenorrhea NIMPUTE=**10** seed=**8675309** out=MI\_Amen noprint;

\* We could use the "by" statement to impute separately for each ID.

\* We don't want to do that here, but it's an option when there is lots of data per person/cluster;

\* by id;

class trt time y;

\* We don’t specify link = logit here but that is what is used.

With multinomial data,

* link = logit will use the cumulative logistic model for ordinal data,
* link = glogit for nominal data;

monotone logistic (y = trt time);

var trt time y;

**run**;

**proc** **print** data = MI\_Amen (obs=**10**);

**run**;

|  |
| --- |
| The SAS System |

| **Obs** | **\_Imputation\_** | **ID** | **TRT** | **TIME** | **Y** | **Ctime** | **prevy** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 1 | 1 | 0 | 1 | 0 | 1 | . |
| **2** | 1 | 1 | 0 | 2 | 0 | 2 | 0 |
| **3** | 1 | 1 | 0 | 3 | 1 | 3 | . |
| **4** | 1 | 1 | 0 | 4 | 1 | 4 | . |
| **5** | 1 | 2 | 0 | 1 | 0 | 1 | . |
| **6** | 1 | 2 | 0 | 2 | 0 | 2 | 0 |
| **7** | 1 | 2 | 0 | 3 | 0 | 3 | . |
| **8** | 1 | 2 | 0 | 4 | 1 | 4 | . |
| **9** | 1 | 3 | 0 | 1 | 0 | 1 | . |
| **10** | 1 | 3 | 0 | 2 | 0 | 2 | 0 |

**Comparing the means of the imputed variables we the means of the observed variables.**

**proc** **means** data=MI\_Amen mean stderr ;

by \_imputation\_ trt time;

var y;

output out=MI\_mn mean=mn\_Y stderr=SE\_Y;

**run** ;

**proc** **sort** data=MI\_mn ;

by trt time \_imputation\_ ;

**run** ;

**proc** **mianalyze** data=MI\_mn ;

by trt time;

modeleffects mn\_Y;

stderr SE\_Y;

ods output parameterestimates=outcombine\_1 ;

**run** ;

**proc** **print** data=outcombine\_1;

**run**;

| **Obs** | **TRT** | **TIME** | **NImpute** | **Parm** | **Estimate** | **StdErr** | **LCLMean** | **UCLMean** | **DF** | **Min** | **Max** | **Theta0** | **tValue** | **Probt** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 0 | 1 | 10 | mn\_Y | 0.185764 | 0.016219 | . | . | . | 0.185764 | 0.185764 | 0 | . | . |
| **2** | 0 | 2 | 10 | mn\_Y | 0.263542 | 0.018882 | 0.2265 | 0.3006 | 3166 | 0.256944 | 0.272569 | 0 | 13.96 | <.0001 |
| **3** | 0 | 3 | 10 | mn\_Y | 0.391667 | 0.023823 | 0.3445 | 0.4388 | 123.2 | 0.368056 | 0.409722 | 0 | 16.44 | <.0001 |
| **4** | 0 | 4 | 10 | mn\_Y | 0.499653 | 0.025585 | 0.4487 | 0.5506 | 79.577 | 0.480903 | 0.519097 | 0 | 19.53 | <.0001 |
| **5** | 1 | 1 | 10 | mn\_Y | 0.205217 | 0.016857 | . | . | . | 0.205217 | 0.205217 | 0 | . | . |
| **6** | 1 | 2 | 10 | mn\_Y | 0.331826 | 0.021274 | 0.2900 | 0.3736 | 418.18 | 0.320000 | 0.344348 | 0 | 15.60 | <.0001 |
| **7** | 1 | 3 | 10 | mn\_Y | 0.491130 | 0.025546 | 0.4403 | 0.5420 | 81.016 | 0.467826 | 0.520000 | 0 | 19.23 | <.0001 |
| **8** | 1 | 4 | 10 | mn\_Y | 0.542783 | 0.025184 | 0.4927 | 0.5928 | 88.573 | 0.518261 | 0.566957 | 0 | 21.55 | <.0001 |

**proc** **means** data = BIOS755.Amenorrhea\_wide mean stderr clm alpha=**0.05**;

var Y1 Y2 Y3 Y4;

by TRT;

**run**;

trt=0

| **Variable** | **Mean** | **Std Error** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Y1** | | **Y2** | | **Y3** | | **Y4** | | |  | | --- | | 0.1857639 | | 0.2620545 | | 0.3887531 | | 0.5013850 | | |  | | --- | | 0.0162189 | | 0.0201560 | | 0.0241332 | | 0.0263522 | | |  | | --- | | 0.1539083 | | 0.2224488 | | 0.3413121 | | 0.4495614 | | |  | | --- | | 0.2176194 | | 0.3016603 | | 0.4361940 | | 0.5532087 | |

trt=1

| **Variable** | **Mean** | **Std Error** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Y1** | | **Y2** | | **Y3** | | **Y4** | | |  | | --- | | 0.2052174 | | 0.3361345 | | 0.4935733 | | 0.5354108 | | |  | | --- | | 0.0168568 | | 0.0216745 | | 0.0253816 | | 0.0265832 | | |  | | --- | | 0.1721088 | | 0.2935446 | | 0.4436707 | | 0.4831289 | | |  | | --- | | 0.2383259 | | 0.3787243 | | 0.5434759 | | 0.5876926 | |

**The second way of imputing is going to use the monotone statement on the wide version of the dataset.**

**proc** **transpose** data=BIOS755.Amenorrhea out=Amenorrhea\_wide prefix=Y;

by id;

id time;

var y;

**run**;

**data** BIOS755.Amenorrhea\_wide;

merge BIOS755.Amenorrhea Amenorrhea\_wide;

by id;

if first.id;

drop y \_NAME\_ \_LABEL\_;

**run**;

**proc** **print** data = BIOS755.Amenorrhea\_wide (obs=**10**);

**run**;

| **Obs** | **ID** | **TRT** | **TIME** | **Y1** | **Y2** | **Y3** | **Y4** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 1 | 0 | 1 | 0 | . | . | . |
| **2** | 2 | 0 | 1 | 0 | . | . | . |
| **3** | 3 | 0 | 1 | 0 | . | . | . |
| **4** | 4 | 0 | 1 | 0 | . | . | . |
| **5** | 5 | 0 | 1 | 0 | . | . | . |
| **6** | 6 | 0 | 1 | 0 | . | . | . |
| **7** | 7 | 0 | 1 | 0 | . | . | . |
| **8** | 8 | 0 | 1 | 0 | . | . | . |
| **9** | 9 | 0 | 1 | 0 | . | . | . |
| **10** | 10 | 0 | 1 | 0 | . | . | . |

**proc** **sort** data = BIOS755.Amenorrhea\_wide;

by TRT;

**run**;

**proc** **mi** data=BIOS755.Amenorrhea\_wide NIMPUTE=**10** seed=**8675309** out=MI\_wide;

class y1 - y4;

monotone logistic (y2 = y1 trt);

monotone logistic (y3 = y2 y1 y1\*y2 trt);

monotone logistic (y4 = y3 y2 y1 y1\*y2 y3\*y2 trt);

var trt y1 - y4;

**run**;

| **Model Information** | |
| --- | --- |
| **Data Set** | BIOS755.AMENORRHEA\_WIDE |
| **Method** | Monotone |
| **Number of Imputations** | 10 |
| **Seed for random number generator** | 8675309 |

| **Monotone Model Specification** | |
| --- | --- |
| **Method** | **Imputed Variables** |
| Logistic Regression | Y2 Y3 Y4 |
| Discriminant Function | Y1 |

| **Missing Data Patterns** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Group** | **TRT** | **Y1** | **Y2** | **Y3** | **Y4** | **Freq** | **Percent** | **Group Means** |
| **TRT** |
| **1** | X | X | X | X | X | 714 | 62.03 | 0.494398 |
| **2** | X | X | X | X | . | 84 | 7.30 | 0.428571 |
| **3** | X | X | X | . | . | 155 | 13.47 | 0.561290 |
| **4** | X | X | . | . | . | 198 | 17.20 | 0.500000 |

**Then, I’ll do the same things I did before. Comparing the means of the imputed variables we the means of the observed variables.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Obs** | **TRT** | **NImpute** | **Parm** | **Estimate** | **StdErr** | **LCLMean** | **UCLMean** | **DF** | **Min** | **Max** | **Theta0** | **tValue** | **Probt** |
| **1** | 0 | 10 | mn\_Y1 | 0.18576 | 0.01622 | . | . | . | 0.18576 | 0.18576 | 0 | . | . |
| **2** | 0 | 10 | mn\_Y2 | 0.26892 | 0.01986 | 0.2299 | 0.30795 | 505.4 | 0.26042 | 0.27951 | 0 | 13.54 | <.0001 |
| **3** | 0 | 10 | mn\_Y3 | 0.39983 | 0.02587 | 0.34815 | 0.45151 | 63.42 | 0.37326 | 0.42361 | 0 | 15.46 | <.0001 |
| **4** | 0 | 10 | mn\_Y4 | 0.51962 | 0.02529 | 0.46936 | 0.56988 | 87.05 | 0.49653 | 0.54167 | 0 | 20.55 | <.0001 |
| **5** | 1 | 10 | mn\_Y1 | 0.20522 | 0.01686 | . | . | . | 0.20522 | 0.20522 | 0 | . | . |
| **6** | 1 | 10 | mn\_Y2 | 0.34661 | 0.02164 | 0.30405 | 0.38917 | 360.9 | 0.33391 | 0.35652 | 0 | 16.01 | <.0001 |
| **7** | 1 | 10 | mn\_Y3 | 0.51478 | 0.02513 | 0.46488 | 0.56469 | 92.81 | 0.49044 | 0.5287 | 0 | 20.48 | <.0001 |
| **8** | 1 | 10 | mn\_Y4 | 0.5673 | 0.02477 | 0.51815 | 0.61646 | 97.76 | 0.54087 | 0.58435 | 0 | 22.9 | <.0001 |

The MEANS Procedure

trt=0

| **Variable** | **Mean** | **Std Dev** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Y1** | | **Y2** | | **Y3** | | **Y4** | | |  | | --- | | 0.1857639 | | 0.2620545 | | 0.3887531 | | 0.5013850 | | |  | | --- | | 0.3892541 | | 0.4402138 | | 0.4880641 | | 0.5006920 | | |  | | --- | | 0.1539083 | | 0.2224488 | | 0.3413121 | | 0.4495614 | | |  | | --- | | 0.2176194 | | 0.3016603 | | 0.4361940 | | 0.5532087 | |

trt=1

| **Variable** | **Mean** | **Std Dev** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Y1** | | **Y2** | | **Y3** | | **Y4** | | |  | | --- | | 0.2052174 | | 0.3361345 | | 0.4935733 | | 0.5354108 | | |  | | --- | | 0.4042120 | | 0.4728825 | | 0.5006026 | | 0.4994524 | | |  | | --- | | 0.1721088 | | 0.2935446 | | 0.4436707 | | 0.4831289 | | |  | | --- | | 0.2383259 | | 0.3787243 | | 0.5434759 | | 0.5876926 | |

**Go from wide to long:**

**data** MI\_wide\_long;

set MI\_wide;

time=**1**;

Y=Y1;

output;

time=**2**;

Y=Y2;

output;

time=**3**;

Y=Y3;

output;

time=**4**;

Y=Y4;

output;

drop Y1-Y4;

**run**;

**run**;

**Now, we’ll analyze the data using proc glimmix by \_Imputation\_:**

**proc** **glimmix** data=MI\_wide\_long method=QUAD(qpoints=**50**);

by \_Imputation\_;

class ID;

model Y = Time trt Time\*Time trt\*Time trt\*Time\*Time / dist=bin solution covb;

random intercept / subject=id;

ods output ParameterEstimates=mixparms CovB=mixcovb;

run;

**The output from the first 2 imputations:**

|  |
| --- |
| The SAS System |

The GLIMMIX Procedure

Imputation Number=1

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MI\_WIDE\_LONG |
| **Response Variable** | Y |
| **Response Distribution** | Binomial |
| **Link Function** | Logit |
| **Variance Function** | Default |
| **Variance Matrix Blocked By** | ID |
| **Estimation Technique** | Maximum Likelihood |
| **Likelihood Approximation** | Gauss-Hermite Quadrature |
| **Degrees of Freedom Method** | Containment |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **ID** | 1151 | 1 2 3 4 5 6 7 8 9 ……. 1151 |

|  |  |
| --- | --- |
| **Number of Observations Read** | 4604 |
| **Number of Observations Used** | 4604 |

| **Dimensions** | |
| --- | --- |
| **G-side Cov. Parameters** | 1 |
| **Columns in X** | 6 |
| **Columns in Z per Subject** | 1 |
| **Subjects (Blocks in V)** | 1151 |
| **Max Obs per Subject** | 4 |

| **Optimization Information** | |
| --- | --- |
| **Optimization Technique** | Dual Quasi-Newton |
| **Parameters in Optimization** | 7 |
| **Lower Boundaries** | 1 |
| **Upper Boundaries** | 0 |
| **Fixed Effects** | Not Profiled |
| **Starting From** | GLM estimates |
| **Quadrature Points** | 50 |

| **Iteration History** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Iteration** | **Restarts** | **Evaluations** | **Objective Function** | **Change** | **Max Gradient** |
| **0** | **0** | 4 | 5076.0941281 | . | 1097.241 |
| **1** | **0** | 4 | 5067.7387036 | 8.35542455 | 201.1249 |
| **2** | **0** | 2 | 5049.0448135 | 18.69389005 | 726.9587 |
| **3** | **0** | 3 | 5029.056989 | 19.98782452 | 192.037 |
| **14** | **0** | 3 | 4963.0965641 | 0.00157926 | 0.169368 |
| **15** | **0** | 3 | 4963.096537 | 0.00002709 | 0.015629 |

|  |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

| **Fit Statistics** | |
| --- | --- |
| **-2 Log Likelihood** | 4963.10 |
| **AIC (smaller is better)** | 4977.10 |
| **AICC (smaller is better)** | 4977.12 |
| **BIC (smaller is better)** | 5012.44 |
| **CAIC (smaller is better)** | 5019.44 |
| **HQIC (smaller is better)** | 4990.44 |

| **Fit Statistics for Conditional Distribution** | |
| --- | --- |
| **-2 log L(Y | r. effects)** | 2950.53 |
| **Pearson Chi-Square** | 2486.32 |
| **Pearson Chi-Square / DF** | 0.54 |

| **Covariance Parameter Estimates** | | | |
| --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Intercept** | **ID** | 4.8233 | 0.4605 |

| **Solutions for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | -3.5029 | 0.3822 | 1149 | -9.16 | <.0001 |
| **TIME** | 0.9244 | 0.3135 | 3449 | 2.95 | 0.0032 |
| **TRT** | -0.6139 | 0.5237 | 3449 | -1.17 | 0.2412 |
| **TIME\*TIME** | -0.00429 | 0.06015 | 3449 | -0.07 | 0.9432 |
| **TIME\*TRT** | 1.0102 | 0.4383 | 3449 | 2.30 | 0.0213 |
| **TIME\*TIME\*TRT** | -0.1831 | 0.08420 | 3449 | -2.17 | 0.0298 |

| **Type III Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **TIME** | 1 | 3449 | 8.70 | 0.0032 |
| **TRT** | 1 | 3449 | 1.37 | 0.2412 |
| **TIME\*TIME** | 1 | 3449 | 0.01 | 0.9432 |
| **TIME\*TRT** | 1 | 3449 | 5.31 | 0.0213 |
| **TIME\*TIME\*TRT** | 1 | 3449 | 4.73 | 0.0298 |

| **Covariance Matrix for Fixed Effects** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Effect** | **Row** | **Col1** | **Col2** | **Col3** | **Col4** | **Col5** | **Col6** |
| **Intercept** | **1** | 0.1461 | -0.1090 | -0.1392 | 0.01924 | 0.1060 | -0.01897 |
| **TIME** | **2** | -0.1090 | 0.09826 | 0.1075 | -0.01850 | -0.09762 | 0.01844 |
| **TRT** | **3** | -0.1392 | 0.1075 | 0.2743 | -0.01929 | -0.2112 | 0.03785 |
| **TIME\*TIME** | **4** | 0.01924 | -0.01850 | -0.01929 | 0.003618 | 0.01852 | -0.00362 |
| **TIME\*TRT** | **5** | 0.1060 | -0.09762 | -0.2112 | 0.01852 | 0.1921 | -0.03627 |
| **TIME\*TIME\*TRT** | **6** | -0.01897 | 0.01844 | 0.03785 | -0.00362 | -0.03627 | 0.007090 |

|  |
| --- |
| The SAS System |

The GLIMMIX Procedure

Imputation Number=2

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MI\_WIDE\_LONG |
| **Response Variable** | Y |
| **Response Distribution** | Binomial |
| **Link Function** | Logit |
| **Variance Function** | Default |
| **Variance Matrix Blocked By** | ID |
| **Estimation Technique** | Maximum Likelihood |
| **Likelihood Approximation** | Gauss-Hermite Quadrature |
| **Degrees of Freedom Method** | Containment |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **ID** | 1151 | 1 2 3 4 5 6…. 1151 |

|  |  |
| --- | --- |
| **Number of Observations Read** | 4604 |
| **Number of Observations Used** | 4604 |

| **Dimensions** | |
| --- | --- |
| **G-side Cov. Parameters** | 1 |
| **Columns in X** | 6 |
| **Columns in Z per Subject** | 1 |
| **Subjects (Blocks in V)** | 1151 |
| **Max Obs per Subject** | 4 |

| **Optimization Information** | |
| --- | --- |
| **Optimization Technique** | Dual Quasi-Newton |
| **Parameters in Optimization** | 7 |
| **Lower Boundaries** | 1 |
| **Upper Boundaries** | 0 |
| **Fixed Effects** | Not Profiled |
| **Starting From** | GLM estimates |
| **Quadrature Points** | 50 |

| **Iteration History** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Iteration** | **Restarts** | **Evaluations** | **Objective Function** | **Change** | **Max Gradient** |
| **0** | **0** | 4 | 5002.670221 | . | 989.6732 |
| **1** | **0** | 4 | 4995.2095204 | 7.46070063 | 208.1282 |
| **10** | **0** | 3 | 4873.2681437 | 0.00026587 | 0.172507 |
| **11** | **0** | 4 | 4873.2677845 | 0.00035921 | 2.552924 |

|  |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

| **Fit Statistics** | |
| --- | --- |
| **-2 Log Likelihood** | 4873.27 |
| **AIC (smaller is better)** | 4887.27 |
| **AICC (smaller is better)** | 4887.29 |
| **BIC (smaller is better)** | 4922.61 |
| **CAIC (smaller is better)** | 4929.61 |
| **HQIC (smaller is better)** | 4900.61 |

| **Fit Statistics for Conditional Distribution** | |
| --- | --- |
| **-2 log L(Y | r. effects)** | 2815.05 |
| **Pearson Chi-Square** | 2376.58 |
| **Pearson Chi-Square / DF** | 0.52 |

| **Covariance Parameter Estimates** | | | |
| --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Intercept** | **ID** | 5.4393 | 0.5186 |

| **Solutions for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | -3.3388 | 0.3893 | 1149 | -8.58 | <.0001 |
| **TIME** | 0.6620 | 0.3198 | 3449 | 2.07 | 0.0385 |
| **TRT** | -1.0295 | 0.5360 | 3449 | -1.92 | 0.0549 |
| **TIME\*TIME** | 0.04142 | 0.06155 | 3449 | 0.67 | 0.5010 |
| **TIME\*TRT** | 1.4343 | 0.4483 | 3449 | 3.20 | 0.0014 |
| **TIME\*TIME\*TRT** | -0.2567 | 0.08612 | 3449 | -2.98 | 0.0029 |

| **Type III Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **TIME** | 1 | 3449 | 4.28 | 0.0385 |
| **TRT** | 1 | 3449 | 3.69 | 0.0549 |
| **TIME\*TIME** | 1 | 3449 | 0.45 | 0.5010 |
| **TIME\*TRT** | 1 | 3449 | 10.24 | 0.0014 |
| **TIME\*TIME\*TRT** | 1 | 3449 | 8.88 | 0.0029 |

| **Covariance Matrix for Fixed Effects** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Effect** | **Row** | **Col1** | **Col2** | **Col3** | **Col4** | **Col5** | **Col6** |
| **Intercept** | **1** | 0.1515 | -0.1124 | -0.1441 | 0.01989 | 0.1091 | -0.01957 |
| **TIME** | **2** | -0.1124 | 0.1023 | 0.1113 | -0.01931 | -0.1018 | 0.01926 |
| **TRT** | **3** | -0.1441 | 0.1113 | 0.2873 | -0.02004 | -0.2203 | 0.03947 |
| **TIME\*TIME** | **4** | 0.01989 | -0.01931 | -0.02004 | 0.003788 | 0.01938 | -0.00379 |
| **TIME\*TRT** | **5** | 0.1091 | -0.1018 | -0.2203 | 0.01938 | 0.2010 | -0.03795 |
| **TIME\*TIME\*TRT** | **6** | -0.01957 | 0.01926 | 0.03947 | -0.00379 | -0.03795 | 0.007417 |

**Now, we’ll pool the imputed analyses into one set of parameter estimates with proc mialalyze:**

**proc** **mianalyze** parms=mixparms covb(effectvar=rowcol)=mixcovb ;

\*The covb is only needed for multi-variate results (i.e., type III tests or contrast statements;

modeleffects Intercept Time trt Time\*Time trt\*Time trt\*Time\*Time;

title 'MIANALYZE Results';

**run**;

|  |
| --- |
| MIANALYZE Results |

The MIANALYZE Procedure

| **Model Information** | |
| --- | --- |
| **PARMS Data Set** | WORK.MIXPARMS |
| **COVB Data Set** | WORK.MIXCOVB |
| **Number of Imputations** | 10 |

| **Variance Information (10 Imputations)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Variance** | | | **DF** | **Relative Increase in Variance** | **Fraction Missing Information** | **Relative Efficiency** |
| **Between** | **Within** | **Total** |
| **Intercept** | 0.023032 | 0.143971 | 0.169306 | 401.93 | 0.175972 | 0.153840 | 0.984849 |
| **Time** | 0.025113 | 0.097413 | 0.125037 | 184.4 | 0.283573 | 0.229240 | 0.977590 |
| **trt** | 0.042740 | 0.272122 | 0.319136 | 414.7 | 0.172769 | 0.151400 | 0.985086 |
| **Time\*Time** | 0.000937 | 0.003597 | 0.004628 | 181.35 | 0.286626 | 0.231205 | 0.977402 |
| **Time\*trt** | 0.058482 | 0.191145 | 0.255475 | 141.94 | 0.336551 | 0.262130 | 0.974457 |
| **Time\*Time\*trt** | 0.002635 | 0.007060 | 0.009959 | 106.25 | 0.410524 | 0.304022 | 0.970495 |

| **Parameter Estimates (10 Imputations)** | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** | **Std Error** | **95% Confidence Limits** | | **DF** | **Minimum** | **Maximum** | **Theta0** | **t for H0: Parameter=Theta0** | **Pr > |t|** |
| **Intercept** | -3.374046 | 0.411468 | -4.18294 | -2.56515 | 401.93 | -3.680824 | -3.104243 | 0 | -8.20 | <.0001 |
| **Time** | 0.821449 | 0.353606 | 0.12382 | 1.51908 | 184.4 | 0.627280 | 1.182889 | 0 | 2.32 | 0.0213 |
| **trt** | -0.769218 | 0.564921 | -1.87968 | 0.34125 | 414.7 | -1.029541 | -0.426314 | 0 | -1.36 | 0.1741 |
| **Time\*Time** | 0.014475 | 0.068028 | -0.11975 | 0.14870 | 181.35 | -0.058471 | 0.043144 | 0 | 0.21 | 0.8317 |
| **Time\*trt** | 1.170986 | 0.505445 | 0.17181 | 2.17016 | 141.94 | 0.757193 | 1.460007 | 0 | 2.32 | 0.0219 |
| **Time\*Time\*trt** | -0.222907 | 0.099794 | -0.42075 | -0.02506 | 106.25 | -0.301403 | -0.134900 | 0 | -2.23 | 0.0276 |

**Comparing with the complete data analysis**

**proc** **glimmix** data=BIOS755.Amenorrhea method=QUAD(qpoints=**50**);

class ID;

model Y = Time trt Time\*Time trt\*Time trt\*Time\*Time / dist=bin solution;

random intercept / subject=id;

run;

| **Model Information** | |
| --- | --- |
| **Data Set** | BIOS755.AMENORRHEA |
| **Response Variable** | Y |
| **Response Distribution** | Binomial |
| **Link Function** | Logit |
| **Variance Function** | Default |
| **Variance Matrix Blocked By** | ID |
| **Estimation Technique** | Maximum Likelihood |
| **Likelihood Approximation** | Gauss-Hermite Quadrature |
| **Degrees of Freedom Method** | Containment |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **ID** | 1151 | 1 …. 1150 1151 |

|  |  |
| --- | --- |
| **Number of Observations Read** | 4604 |
| **Number of Observations Used** | 3616 |

| **Dimensions** | |
| --- | --- |
| **G-side Cov. Parameters** | 1 |
| **Columns in X** | 6 |
| **Columns in Z per Subject** | 1 |
| **Subjects (Blocks in V)** | 1151 |
| **Max Obs per Subject** | 4 |

| **Optimization Information** | |
| --- | --- |
| **Optimization Technique** | Dual Quasi-Newton |
| **Parameters in Optimization** | 7 |
| **Lower Boundaries** | 1 |
| **Upper Boundaries** | 0 |
| **Fixed Effects** | Not Profiled |
| **Starting From** | GLM estimates |
| **Quadrature Points** | 50 |

| **Iteration History** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Iteration** | **Restarts** | **Evaluations** | **Objective Function** | **Change** | **Max Gradient** |
| **0** | **0** | 4 | 3961.1107127 | . | 754.8834 |
| **1** | **0** | 4 | 3954.5188083 | 6.59190441 | 170.3103 |
| **2** | **0** | 2 | 3935.2222761 | 19.29653216 | 501.6977 |
| **3** | **0** | 3 | 3916.3366324 | 18.88564370 | 125.6327 |
| **4** | **0** | 2 | 3895.3759089 | 20.96072351 | 37.89303 |
| **5** | **0** | 4 | 3875.3319661 | 20.04394273 | 102.5231 |
| **6** | **0** | 3 | 3867.4836672 | 7.84829892 | 86.57043 |
| **7** | **0** | 3 | 3867.1121884 | 0.37147880 | 3.268637 |
| **8** | **0** | 3 | 3866.9314898 | 0.18069861 | 35.01058 |
| **9** | **0** | 3 | 3866.8928583 | 0.03863151 | 3.791026 |
| **10** | **0** | 3 | 3866.8913504 | 0.00150794 | 0.307824 |
| **11** | **0** | 2 | 3866.8899544 | 0.00139600 | 1.891312 |
| **12** | **0** | 2 | 3866.8877416 | 0.00221279 | 0.181709 |
| **13** | **0** | 3 | 3866.8874462 | 0.00029532 | 0.43449 |
| **14** | **0** | 3 | 3866.8874348 | 0.00001144 | 0.012497 |

|  |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

| **Fit Statistics** | |
| --- | --- |
| **-2 Log Likelihood** | 3866.89 |
| **AIC (smaller is better)** | 3880.89 |
| **AICC (smaller is better)** | 3880.92 |
| **BIC (smaller is better)** | 3916.23 |
| **CAIC (smaller is better)** | 3923.23 |
| **HQIC (smaller is better)** | 3894.23 |

| **Fit Statistics for Conditional Distribution** | |
| --- | --- |
| **-2 log L(Y | r. effects)** | 2169.87 |
| **Pearson Chi-Square** | 1741.85 |
| **Pearson Chi-Square / DF** | 0.48 |

| **Covariance Parameter Estimates** | | | |
| --- | --- | --- | --- |
| **Cov Parm** | **Subject** | **Estimate** | **Standard Error** |
| **Intercept** | ID | 5.0798 | 0.5855 |

| **Solutions for Fixed Effects** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Effect** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | -3.3943 | 0.4146 | 1149 | -8.19 | <.0001 |
| **TIME** | 0.7978 | 0.3551 | 2461 | 2.25 | 0.0248 |
| **TRT** | -0.8049 | 0.5636 | 2461 | -1.43 | 0.1534 |
| **TIME\*TIME** | 0.01966 | 0.06968 | 2461 | 0.28 | 0.7779 |
| **TIME\*TRT** | 1.2194 | 0.4978 | 2461 | 2.45 | 0.0144 |
| **TIME\*TIME\*TRT** | -0.2296 | 0.09770 | 2461 | -2.35 | 0.0188 |

| **Type III Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **TIME** | 1 | 2461 | 5.05 | 0.0248 |
| **TRT** | 1 | 2461 | 2.04 | 0.1534 |
| **TIME\*TIME** | 1 | 2461 | 0.08 | 0.7779 |
| **TIME\*TRT** | 1 | 2461 | 6.00 | 0.0144 |
| **TIME\*TIME\*TRT** | 1 | 2461 | 5.52 | 0.0188 |

**Below, we’ll use a weighted GEE to analyze the data:**

**First we create a version of the data with the lagged outcome and a second time variable.**

**data** BIOS755.Amenorrhea;

set BIOS755.Amenorrhea;

Ctime = time;

prevy = lag(y);

**run**;

**Below the missmodel command is what estimates the weights.**

**For the weight model we are using categorical time (profile analysis), for the actual model we’re using continuous time. This is why we needed two time variables.**

ods graphics on;

**proc** **gee** data=BIOS755.Amenorrhea desc plots=histogram;

class ID Ctime;

missmodel Ctime Prevy trt trt\*Prevy / type=obslevel;

model Y = Time trt Time\*Time trt\*Time trt\*Time\*Time / dist=bin;

repeated subject=ID / within=Ctime corr=cs;

**run**;

This runs the following model:

which is used to estimate

| **Model Information** | | | |
| --- | --- | --- | --- |
| **Data Set** | BIOS755.AMENORRHEA |  | y |
| **Distribution** | Binomial |  |  |
| **Link Function** | Logit |  |  |
| **Dependent Variable** | Y |  |  |

|  |  |
| --- | --- |
| **Number of Observations Read** | 4604 |
| **Number of Observations Used** | 3616 |
| **Number of Events** | 1231 |
| **Number of Trials** | 3616 |
| **Number of Missing Values** | 988 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **ID** | 1151 | 1 2 3 4 5 6 7 8 9 10 …. 77 78 79 80 81 82 83 84 85 86 87 ... |
| **Ctime** | 4 | 1 2 3 4 |

| **Response Profile** | | |
| --- | --- | --- |
| **Ordered Value** | **Y** | **Total Frequency** |
| **1** | 1 | 1231 |
| **2** | 0 | 2385 |

|  |
| --- |
| ***PROC GEE is modeling the probability that Y='1'.*** |

| **GEE Model Information** | |
| --- | --- |
| **Correlation Structure** | Exchangeable |
| **Within-Subject Effect** | Ctime (4 levels) |
| **Subject Effect** | ID (1151 levels) |
| **Number of Clusters** | 1151 |
| **Clusters With Missing Values** | 437 |
| **Correlation Matrix Dimension** | 4 |
| **Maximum Cluster Size** | 4 |
| **Minimum Cluster Size** | 1 |

| **Missing Data Patterns** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Group** | **Occasion** | | | | **Freq** | **Percent** |
| **1** | **2** | **3** | **4** |
| **1** | X | X | X | X | 714 | 62.03 |
| **2** | X | X | X | . | 84 | 7.30 |
| **3** | X | X | . | . | 155 | 13.47 |
| **4** | X | . | . | . | 198 | 17.20 |

| **Exchangeable Working Correlation** | |
| --- | --- |
| **Correlation** | 0.3949 |

| **Parameter Estimates for Response Model** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **with Empirical Standard Error Estimates** | | | | | | |
| **Parameter** | **Estimate** | **Standard Error** | **95% Confidence Limits** | | **Z** | **Pr > |Z|** |
| **Intercept** | -2.0381 | 0.2484 | -2.5251 | -1.5512 | -8.20 | <.0001 |
| **TIME** | 0.5453 | 0.2117 | 0.1303 | 0.9603 | 2.58 | 0.0100 |
| **TRT** | -0.4296 | 0.3540 | -1.1234 | 0.2643 | -1.21 | 0.2250 |
| **TIME\*TIME** | -0.0037 | 0.0405 | -0.0831 | 0.0757 | -0.09 | 0.9275 |
| **TRT\*TIME** | 0.6621 | 0.3021 | 0.0700 | 1.2542 | 2.19 | 0.0284 |
| **TRT\*TIME\*TIME** | -0.1264 | 0.0577 | -0.2395 | -0.0134 | -2.19 | 0.0284 |

| **Parameter Estimates for Missingness Model** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** |  | **Estimate** | **Standard Error** | **95% Confidence Limits** | | **Z** | **Pr > |Z|** |
| **Intercept** |  | 2.3967 | 0.1438 | 2.1149 | 2.6785 | 16.67 | <.0001 |
| **Ctime** | 1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| **Ctime** | 2 | -0.7286 | 0.1439 | -1.0106 | -0.4466 | -5.06 | <.0001 |
| **Ctime** | 3 | -0.5919 | 0.1469 | -0.8798 | -0.3040 | -4.03 | <.0001 |
| **Ctime** | 4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| **prevy** |  | -0.4514 | 0.1619 | -0.7687 | -0.1341 | -2.79 | 0.0053 |
| **TRT** |  | 0.0680 | 0.1313 | -0.1893 | 0.3253 | 0.52 | 0.6046 |
| **prevy\*TRT** |  | -0.2381 | 0.2196 | -0.6685 | 0.1923 | -1.08 | 0.2782 |



**Comparing with the complete data analysis**

**proc** **gee** data=BIOS755.Amenorrhea desc plots=histogram;

class ID Ctime;

model Y = Time trt Time\*Time trt\*Time trt\*Time\*Time / dist=bin;

repeated subject=ID / within=Ctime corr=cs;

**run**;

| **Model Information** | | | |
| --- | --- | --- | --- |
| **Data Set** | BIOS755.AMENORRHEA |  | y |
| **Distribution** | Binomial |  |  |
| **Link Function** | Logit |  |  |
| **Dependent Variable** | Y |  |  |

|  |  |
| --- | --- |
| **Number of Observations Read** | 4604 |
| **Number of Observations Used** | 3616 |
| **Number of Events** | 1231 |
| **Number of Trials** | 3616 |
| **Number of Missing Values** | 988 |

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **ID** | 1151 | 1 2 3 4 ….. 75 76 77 78 79 80 81 82 83 84 85 86 87 ... |
| **Ctime** | 4 | 1 2 3 4 |

| **Response Profile** | | |
| --- | --- | --- |
| **Ordered Value** | **Y** | **Total Frequency** |
| **1** | 1 | 1231 |
| **2** | 0 | 2385 |

|  |
| --- |
| ***PROC GEE is modeling the probability that Y='1'.*** |

| **GEE Model Information** | |
| --- | --- |
| **Correlation Structure** | Exchangeable |
| **Within-Subject Effect** | Ctime (4 levels) |
| **Subject Effect** | ID (1151 levels) |
| **Number of Clusters** | 1151 |
| **Clusters With Missing Values** | 437 |
| **Correlation Matrix Dimension** | 4 |
| **Maximum Cluster Size** | 4 |
| **Minimum Cluster Size** | 1 |

| **Exchangeable Working Correlation** | |
| --- | --- |
| **Correlation** | 0.3626 |

| **GEE Fit Criteria** | |
| --- | --- |
| QIC | 4384.4585 |
| QICu | 4382.9860 |

| **Parameter Estimates for Response Model** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **with Empirical Standard Error Estimates** | | | | | | |
| **Parameter** | **Estimate** | **Standard Error** | **95% Confidence Limits** | | **Z** | **Pr > |Z|** |
| **Intercept** | -2.0022 | 0.2466 | -2.4856 | -1.5187 | -8.12 | <.0001 |
| **TIME** | 0.5097 | 0.2095 | 0.0990 | 0.9205 | 2.43 | 0.0150 |
| **TRT** | -0.4516 | 0.3514 | -1.1403 | 0.2371 | -1.29 | 0.1987 |
| **TIME\*TIME** | 0.0014 | 0.0401 | -0.0773 | 0.0801 | 0.03 | 0.9724 |
| **TIME\*TRT** | 0.6902 | 0.2988 | 0.1045 | 1.2760 | 2.31 | 0.0209 |
| **TIME\*TIME\*TRT** | -0.1299 | 0.0571 | -0.2419 | -0.0180 | -2.27 | 0.0229 |