# Dataset Integrity Check (DSIC) for the HEALTHY Data Files 

Reference paper:
NEJM 363(5) [2010 Jul 29]:443-53.

HEALTHY is a cluster randomized trial designed to investigate the effectiveness of an integrated intervention in middle schools in the reduction of risk factors for type 2 diabetes. The trial was conducted at 7 field centers in 42 middle schools, randomly assigned to intervention or control. Student recruitment and baseline data collection took place during the first semester of 6th grade (2006); the intervention was implemented the following semester (2007) and continued through $8^{\text {th }}$ grade (2007-2009). The intervention, designed to impact the environment and lifestyle choices of middle school children, was implemented at the school-level. As a partial check of the integrity of the HEALTHY baseline survey dataset archived in the NIDDK data repository, a dataset integrity check (DSIC) was performed to verify that selected published results from the HEALTHY study can be reproduced using the archived dataset. The DSIC consists of a small number of analyses performed to duplicate published results reported by the HEALTHY Study Group [1] in NEJM in July 2010. Results of the DSIC are described below.

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected on a first exercise in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. We do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless staff of the NIDDK Repository suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document in the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

Archived Dataset Contents. The DCC submitted ten $S A S$ transport data files representing the raw data collected from 8 data collection forms, plus files with laboratory values and school randomization assignments. Electronic files of 7 of the 8 data collection forms also were submitted, with fields populated with the corresponding $S A S$ variable names. (The nutrition questionnaire was not submitted due to copyright issues.) The numerical value for check-box style categorical variables was noted to the lower right of the check-boxes on the collection forms.

Dataset deidentification. To prevent identification of participants in the study, the archived datasets were subjected to masking prior to submission to the Repository. All personal identifiers and dates were removed, and school ID was replaced by a randomly generated 2-digit masked ID number. Age at both baseline and end of study was removed. (The distribution of participant age
is described in the accompanying Repository Documentation provided by the DCC, Section 2.3.1). Variables with fine strata that might identify a particular individual were collapsed into wider groupings; e.g., race/ethnicity was coded as "White", "Black", "Hispanic", and "Other"; family history of diabetes was recoded as "first degree" if mother, father or sibling was reported having diabetes and as "second degree" if grandmother, grandfather, or aunt/uncle was reported having diabetes. The highest grade completed in school by the head of household was collapsed into "high school diploma or less", versus "some college or more".

Highest and lowest measures in anthropomorphic data were also collapsed to protect participant identity. Relevant to this analysis, participants with a BMI z-score of at most -0.50 were grouped together, as were those with a BMI z-score of at least 2.10. Also, participants with a waist circumference of at most 57 cm were grouped together, as were those with a waist circumference of at least 97 cm .

Due to deidentification procedures implemented in archived data, results of archived data analyses are expected to differ somewhat from published results. However, we would expect the conclusions of such data analysis to remain the same as what was previously published.

DSIC Analysis Methods. A portion of published results was replicated to ensure integrity of archived datasets.

First, school randomization assignment was matched to each participant via masked school ID number. Next, sixth grade level ('baseline') data were merged with eighth grade ('followup') data, by masked participant ID number, to produce the primary analysis cohort.

Next, distributions of gender, race/ethnicity, education of head of household, and family history of diabetes were calculated and compared to published breakdowns. Outcomes BMI $\geq 85^{\text {th }}$ percentile, $\mathrm{BMI} \geq 95^{\text {th }}$ percentile, fasting insulin $\geq 30 \mathrm{U} / \mathrm{ml}$, and fasting glucose $\geq 100 \mathrm{mg} / \mathrm{dl}$ were calculated for students in control and intervention groups, at baseline ( $6^{\text {th }}$ grade) - and at followup ( $8^{\text {th }}$ grade). Prevalences of outcomes were compared to published numbers. Outcome change scores, from baseline to followup, were calculated for control vs. interventions groups; change scores were compared to published numbers.

Generalized Linear Mixed Modeling (GLMM) was used to assess for treatment differences in bilevel outcomes, adjusting for school as a random effect. Odds ratios and $p$-values were compared to published numbers.

As previously described, participants with extreme values of BMI z-score or waist circumference were collapsed into broad groups. To retain these participants in the DSIC analysis, those coded as having a BMI $z$-score of -0.50 or less were assigned a $z$-score of exactly -0.50 ; those with BMI z-score of 2.10 or more were assigned a z-score of 2.10 . Likewise, participants with a waist circumference of 57 cm or less were assigned a waist circumference of exactly 57 cm ; those with a waist circumference of 97 or more were assigned a waist circumference of 97 cm . Means and standard deviations then were calculated for continuous outcomes: BMI z score, waist circumference, in addition to fasting insulin and fasting glucose. This was done for control versus intervention groups, at baseline and at followup. Mean change scores from baseline to
followup were also calculated. Mean scores of continuous outcomes, and mean change scores, were compared to published results. Mixed modeling was used to assess for treatment differences in continuous change scores, adjusting for school as a random effect. As indicated in the publication, fasting insulin was log-transformed prior to analysis, due to skewness in the data. $P$-values derived from mixed modeling were compared to published results.

All statistical analyses were conducted using SAS version 9.2 (Cary, NC).

DSIC Results: Demographics. Merging of raw datasets as described produced an analysis dataset with 4603 participants in 42 schools. These sample sizes match those of the publication. Distributions of gender, race/ethnicity, education of head of household, and family history of diabetes closely matched published breakdowns; any difference was in the decimal points, easily attributable to rounding. As described in the methods, age was not included in archived data, and could not be analyzed. [Table 1].

Table 1 (continues next page). Baseline Characteristics of the Students: Archived vs. Published Results
(published results extracted from Table 1 in Healthy Study Group, et al., NEJM 363(5), p. 446)

|  | Published |  |
| :--- | :---: | :---: |
|  | Prehived |  |
| No. of students (\% within <br> group) | $4603(100)$ | $4603(100)$ |
| Age (yr) | $11.3+0.6$ | n.a. |
| Male sex (\%) | 47.3 | 47.3 |
| Race or ethnic group (\%) |  |  |
| Hispanic | 54.2 | 54.2 |
| Black | 18.0 | 18.0 |
| White | 19.3 | 19.3 |
| Other | 8.5 | 8.5 |
| Highest education level attained |  |  |
| by head of household (\%)* |  |  |
| HS Diploma or less | 51.7 | 51.7 |
| $\quad$ More than HS Diploma | 48.3 | 48.3 |
| Family history of diabetes (\%)** | 17.6 | 17.6 |

n.a. $=$ Not Available in archived data

Table 1, continued. Baseline Characteristics of the Students: Archived vs. Published Results (published results extracted from Table 1 in Healthy Study Group, et al., NEJM 363(5), p. 446)

## Intervention Group

## Control Group

|  | Total |  | BMI <85th percentile |  | $\frac{\text { BMI }>=85 t h}{\text { percentile }}$ |  | Total |  | BMI <85th percentile |  | $\frac{\text { BMI }>=85 \mathrm{th}}{\text { percentile }}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Published | Archived | Published | Archived | Published | Archived | Published | Archived | Published | Archived | Published | Archived |
| No. of students (\% within group) | 2307 (100) | 2307 (100) | $\begin{aligned} & \hline 1147 \\ & (49.7) \\ & \hline \end{aligned}$ | $\begin{gathered} 1147 \\ (49.7) \\ \hline \end{gathered}$ | $\begin{aligned} & \hline 1160 \\ & (50.3) \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 1160 \\ & (50.3) \\ & \hline \end{aligned}$ | 2296 (100) | $\begin{aligned} & \hline 2296 \\ & (100) \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 1164 \\ & (50.7) \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 1164 \\ & (50.7) \\ & \hline \end{aligned}$ | $\begin{aligned} & 1132 \\ & (49.3) \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 1132 \\ & (49.3) \\ & \hline \end{aligned}$ |
| Age (yr) | $11.3 \pm 0.5$ | n.a. | $11.3 \pm 0.5$ | n.a. | $11.2 \pm 0.5$ | ก.a. | $11.3 \pm 0.6$ | n.a. | $11.3 \pm 0.6$ | n.a. | $11.3 \pm 0.6$ | n.a. |
| Male sex (\%) | 47.4 | 47.4 | 44.2 | 44.2 | 50.5 | 50.5 | 47.1 | 47.1 | 44.0 | 44.0 | 50.4 | 50.4 |
| Race or ethnic group (\%) |  |  |  |  |  |  |  |  |  |  |  |  |
| Hispanic | 54.8 | 54.8 | 51.3 | 51.3 | 58.4 | 58.4 | 53.5 | 53.5 | 50.3 | 50.3 | 56.7 | 56.7 |
| Black | 20.3 | 20.3 | 22.2 | 22.2 | 18.4 | 18.4 | 15.7 | 15.7 | 15.3 | 15.3 | 16.2 | 16.2 |
| White | 17.1 | 17.0 | 18.6 | 18.6 | 15.5 | 15.5 | 21.6 | 21.6 | 24.4 | 24.4 | 18.6 | 18.6 |
| Other | 7.8 | 7.8 | 7.9 | 7.9 | 7.7 | 7.7 | 9.2 | 9.2 | 10.0 | 10.0 | 8.5 | 8.5 |
| Highest education level attained by head of household (\%)* |  |  |  |  |  |  |  |  |  |  |  |  |
| HS Diploma or less | 51.8 | 51.8 | 49.7 | 49.6 | 53.9 | 53.9 | 51.6 | 51.5 | 48.2 | 48.2 | 54.9 | 54.9 |
| More than HS | 48.2 | 48.2 | 50.3 | 50.4 | 46.1 | 46.1 | 48.4 | 48.5 | 51.8 | 51.8 | 45.1 | 45.1 |
| Family history of diabetes (\%)** | 17.1 | 17.1 | 13.0 | 13.0 | 20.9 | 20.9 | 18.1 | 18.1 | 12.4 | 12.4 | 24.2 | 24.2 |

n.a. Not Available in archived data

* In archived data, responses are collapsed into two categories

In archived data, the variable is 'from first degree relative'

DSIC Results: Analysis of Outcomes. Prevalences of bi-level outcomes, i.e., $\mathrm{BMI} \geq 85^{\text {th }}$ percentile, BMI $\geq 95^{\text {th }}$ percentile, fasting insulin $\geq 30 \mathrm{U} / \mathrm{ml}$, and fasting glucose $\geq 100 \mathrm{mg} / \mathrm{dl}$, were calculated for students in control and intervention groups, at baseline ( $6^{\text {th }}$ grade) and at followup ( 8 th grade). Change scores (from baseline to followup) were also calculated by treatment group. Calculated prevalences of bi-level outcomes were similar to published results. Differences were in the decimal places, easily attributable to rounding. Outcome-specific odds ratios and $p$-values for intervention versus control groups were approximately similar in archived versus published results. Differences in odds ratios and $p$-values could be attributed to differences in the application of GLMM methodology (e.g., method used to calculate the variance/covariance ratio, etc.), which were not specified in the publication. [Table 2]

Means and standard deviations were calculated for continuous outcomes, i.e., BMI z-score, waist circumference, fasting insulin, and fasting glucose, by control and intervention groups, at baseline and at followup. Mean change scores were also calculated. Slight differences between published and archived results in continuous outcomes were found. Variances tended to be smaller in archived results compared to those that were published. This is expected, due to the recoding of extreme anthropomorphic measures to a constant in archived data. Mean change scores also differed slightly between published and archived results; the largest discrepancy being the mean change in waist circumference among intervention students ( 3.5 cm in archived results, versus 4.6 cm in published results). Recoding/collapsing of extreme values in the archived dataset may also help explain these differences. $P$-values for differences in treatment effects differed between archived and published results; the differences in variances may help explain this. The discrepancy in $p$-values for the change in waist circumference was $\mathrm{p}=0.07$ (published) versus $\mathrm{p}=0.13$ (archived); for the change in BMI zscore, the discrepancy was $\mathrm{p}=0.04$ (published) versus $\mathrm{p}=0.08$ (archived). [Table 2]

Conclusion. With the replication of selected results, the analysis of archived data closely matches published results, allowing for rounding error and variations expected from data masking. We are confident there were no errors in the transmission of archived datasets from the DCC to the Repository. Approved data users are advised to interpret results of analysis of outcomes with caution (especially that of continuous data outcomes), due to the data masking procedures that were implemented in archived data.

Table 2. Baseline and End-of-Study Assessments of Students in Intervention and Control Schools: Archived vs. Published Results (published results extracted from Table 2 in Healthy Study Group, et al., NEJM 363(5), p. 448)


## References

[1] HEALTHY Study Group, Foster GD, Linder B, Baranowski T, Cooper DM, Goldberg L, Harrell JS, Kaufman F, Marcus MD, Treviño RP, Hirst K. A school-based intervention for diabetes risk reduction. N Engl J Med. 363(5) [2010 Jul 29]: 443-53. Epub 2010 Jun 27.

## Appendices

[1] Full Text of Healthy Study Group, et al., NEJM 363(5), provided to approved data requestors.
[2] SAS version 9.2 Log for programming code submitted for the replication of results in Healthy Study Group, et al., NEJM 363(5)
[3] SAS version 9.2 Output for programming code submitted for the replication of results in Healthy Study Group, et al., NEJM 363(5)

## Attachment 1

## "The full text of the article referenced will be provided to approved data requestors along with the data archived."

[^0]
## Attachment 2

SAS version 9.2 Log for programming code submitted for the replication of results in Tables 1 and 2 of Healthy Study Group, et al., NEJM 363(5)

NOTE: Copyright (c) 2002-2008 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.2 (TS2M2)
Licensed to RTI INTL MAIN, Site 70006746.
NOTE: This session is executing on the XP_PRO platform.
$\begin{array}{lrl}\text { NOTE: } & \text { SAS initialization used: } \\ \text { real time } & 3.36 & \text { seconds } \\ \text { cpu time } & 0.90 \text { seconds }\end{array}$
1 options ps=55 ls=78 nonumber formchar='|----|+\---+=|-^<>*' mprint orientation=portrait
1 ! ;
2
7 * comparison study paper: NEJM 363(5) [2010 Jul 29]:443-53 *
8 * Programmed by: S. Tan


10
11 libname in 'C:\Documents and Settings\stan\My
11 ! Documents \DATA\NIDDK\HEALTHY $\backslash$ Official_Archive\HEALTHY_data';
NOTE: Libref IN was successfully assigned as follows:
Engine: V9
Physical Name: C:\Documents and Settings\stan $\backslash M y$
Documents \DATA\NIDDK $\backslash H E A L T H Y \backslash O f f i c i a l \_A r c h i v e \backslash H E A L T H Y \_d a t a ~$
12
13 proc format;
13 ! value bmige85f $1=' \mathrm{BMI}$ ge 85' 0='BMI lt 85';
NOTE: Format BMIGE85F has been output.
14 value bmige95f 1='BMI ge 95' 0='BMI lt 95';
NOTE: Format BMIGE95F has been output.
15

NOTE: PROCEDURE FORMAT used (Total process time):
real time 0.71 seconds
cpu time 0.04 seconds

16 data st2fam; set in.st2;
NOTE: Data file IN.ST2.DATA is in a format that is native to another host, or the file encoding
does not match the session encoding. Cross Environment Data Access will be
used, which
might require additional CPU resources and might reduce performance.
NOTE: There were 11218 observations read from the data set IN.ST2.
NOTE: The data set WORK.ST2FAM has 11218 observations and 7 variables.
NOTE: DATA statement used (Total process time):
real time 0.04 seconds
cpu time 0.01 seconds

17 data st3screen; set in.st3;
NOTE: Data file IN.ST3.DATA is in a format that is native to another host, or the file encoding
does not match the session encoding. Cross Environment Data Access will be
used, which
might require additional CPU resources and might reduce performance.
NOTE: There were 17896 observations read from the data set IN.ST3.
NOTE: The data set WORK.ST3SCREEN has 17896 observations and 27 variables.
NOTE: DATA statement used (Total process time):
real time 0.06 seconds
cpu time 0.04 seconds

18 data labdata; set in.cbl;
NOTE: Data file IN.CBL.DATA is in a format that is native to another host, or the file encoding does not match the session encoding. Cross Environment Data Access will be used, which
might require additional CPU resources and might reduce performance.
NOTE: There were 12360 observations read from the data set IN.CBL.
NOTE: The data set WORK.LABDATA has 12360 observations and 11 variables.
NOTE: DATA statement used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds

19 data rx; set in.rx; * school randomization assignment: has studentid variable *;
NOTE: Data file IN.RX.DATA is in a format that is native to another host, or the file encoding
does not match the session encoding. Cross Environment Data Access will be
used, which
might require additional CPU resources and might reduce performance.
20

NOTE: There were 42 observations read from the data set IN.RX.
NOTE: The data set WORK.RX has 42 observations and 2 variables.
NOTE: DATA statement used (Total process time):
real time 0.01 seconds
cpu time 0.01 seconds

21
proc sort data=st2fam; by schoolid;
NOTE: There were 11218 observations read from the data set WORK.ST2FAM.
NOTE: The data set WORK.ST2FAM has 11218 observations and 7 variables.
NOTE: PROCEDURE SORT used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds

```
proc sort data=st3screen; by schoolid;
```

NOTE: There were 17896 observations read from the data set WORK.ST3SCREEN.
NOTE: The data set WORK.ST3SCREEN has 17896 observations and 27 variables.
NOTE: PROCEDURE SORT used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds

```
proc sort data=labdata; by schoolid;
```

NOTE: There were 12360 observations read from the data set WORK.LABDATA. NOTE: The data set WORK.LABDATA has 12360 observations and 11 variables. NOTE: PROCEDURE SORT used (Total process time): real time 0.03 seconds cpu time 0.01 seconds

24
25
26
NOTE: There were 42 observations read from the data set WORK.RX. NOTE: The data set WORK.RX has 42 observations and 2 variables.
NOTE: PROCEDURE SORT used (Total process time):
real time 0.06 seconds
cpu time 0.01 seconds

27
data st2fam; merge st2fam rx; by schoolid; run;
NOTE: There were 11218 observations read from the data set WORK.ST2FAM.
NOTE: There were 42 observations read from the data set WORK.RX.
NOTE: The data set WORK.ST2FAM has 11218 observations and 8 variables.
NOTE: DATA statement used (Total process time):
real time 0.01 seconds
cpu time 0.01 seconds

28
data st3screen; merge st3screen rx; by schoolid; run;
NOTE: There were 17896 observations read from the data set WORK.ST3SCREEN.
NOTE: There were 42 observations read from the data set WORK.RX.
NOTE: The data set WORK.ST3SCREEN has 17896 observations and 28 variables.
NOTE: DATA statement used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds

29
data labdata; merge labdata rx; by schoolid; run;
NOTE: There were 12360 observations read from the data set WORK.LABDATA. NOTE: There were 42 observations read from the data set WORK.RX.
NOTE: The data set WORK.LABDATA has 12360 observations and 12 variables. NOTE: DATA statement used (Total process time):
real time 0.01 seconds
cpu time 0.01 seconds

```
proc sort data=st2fam; by studentid;
```

NOTE: There were 11218 observations read from the data set WORK.ST2FAM. NOTE: The data set WORK.ST2FAM has 11218 observations and 8 variables. NOTE: PROCEDURE SORT used (Total process time):

$$
\text { real time } \quad 0.01 \text { seconds }
$$

$$
\text { cpu time } \quad 0.01 \text { seconds }
$$

32
33
34
35

> data g6_st2 97 st2 g8_st2; set st2fam; if grade 6 then output 96 st2; else if grade=7 then output 97 st2; else if grade=8 then output 98 _st2;

NOTE: There were 11218 observations read from the data set WORK.ST2FAM. NOTE: The data set WORK.G6_ST2 has 6288 observations and 8 variables. NOTE: The data set WORK.G7_ST2 has 0 observations and 8 variables. NOTE: The data set WORK.G8_ST2 has 4930 observations and 8 variables. NOTE: DATA statement used (Total process time): real time 0.03 seconds cpu time 0.01 seconds

36 proc sort data=st3screen; by studentid;

NOTE: There were 17896 observations read from the data set WORK.ST3SCREEN. NOTE: The data set WORK.ST3SCREEN has 17896 observations and 28 variables. NOTE: PROCEDURE SORT used (Total process time): real time 0.04 seconds
cpu time 0.03 seconds

37
38
39
40

```
data g6_st3 g7_st3 g8_st3; set st3screen;
    if grade=6 then output g6_st3;
    else if grade=7 then output g7_st3;
    else if grade=8 then output g8_st3;
```

NOTE: There were 17896 observations read from the data set WORK.ST3SCREEN. NOTE: The data set WORK.G6 ST3 has 6358 observations and 28 variables. NOTE: The data set WORK.G7 ST3 has 5225 observations and 28 variables. NOTE: The data set WORK.G8_ST3 has 6313 observations and 28 variables. NOTE: DATA statement used (Total process time):
real time 0.04 seconds
cpu time 0.03 seconds

41
proc sort data=labdata; by studentid;
NOTE: There were 12360 observations read from the data set WORK.LABDATA. NOTE: The data set WORK.LABDATA has 12360 observations and 12 variables. NOTE: PROCEDURE SORT used (Total process time):
real time 0.01 seconds
cpu time $\quad 0.01$ seconds

42
43
44
45
46

```
data g6_lab g7_lab g8_lab; set labdata;
```

    if grade=6 then output g6_lab;
    else if grade \(=7\) then outpūt 97 lab;
    else if grade=8 then output g8_lab;
    NOTE: There were 12360 observations read from the data set WORK.LABDATA. NOTE: The data set WORK.G6_LAB has 6173 observations and 12 variables. NOTE: The data set WORK.G7_LAB has 0 observations and 12 variables.
NOTE: The data set WORK.G8_LAB has 6187 observations and 12 variables.
NOTE: DATA statement used (Total process time):
real time 0.03 seconds
cpu time 0.01 seconds

47 data baseline_f; merge g6_st3(in=in1 rename=(bmipct=bmipct_i bmiz=bmiz_i
47 ! waist=waist_i)
g6_st2 g6_lab(rename=(glucose=glucose_i insulin=insulin_i))
g8_st3(in=in2 keep=studentid bmipct b̄̄iz waist rename=( $\overline{\mathrm{b}} \mathrm{mipct}=\mathrm{bmipct} \mathrm{f}$
49 bmiz=bmiz_f
49 ! waist=waist_f))
50 g8_lab(keēp=studentid glucose insulin rename=(glucose=glucose_f
insulin=insulin_f));
51 by studentid;
52 if in1 and in2; /* the final analysis dataset has students with both
grade 6 and
52 ! grade 8 data */
53 /* n=4603 */
54 * deidentification of data described on page 3 of documentation notes
provided by
54 ! the D.C.C. *;
55
$56 \quad * * * * * * * * * * * * * * *$ Coding of Baseline Measures (6th grade) *;
57 if $30<$ bmipct_i<85 or bmipct_i=1 then bmige85_i=0;
58 else if bmipct_i>=85 or bmipct_i=3 then bmige85_i=1;
59 if $30<$ bmipct_i<95 or bmipct_i=1 then bmige95_i=0;
60 else if bmípct_i>=95 or bmipct_i=3 then bmíge95_i=1;
61 format bmige85_i bmige85f. bmige95_i bmige95f.;
62
63 * recode values >upper limit, or <lower limit, to a constant equal to
exactly that
63 ! limit *;
64 * see forms for values of upper and lower limits *;
65 bmizscore_i=bmiz_i; if bmiz_i=3 then bmizscore_i=2.1; * upper limit
recoded *;
66 if bmiz_i=1 then bmizscore_i=-0.5; * lower limit recoded *;
67 waistcm_i=waist_i; if waist_ $\bar{i}=3$ then waistcm_i=97;* upper limit recoded
*;
68 if waist_i=1 then waistcm_i=57; * lower limit recoded *;
69
70
71 else if $\overline{0}<g l u c o s e \_i<100$ then fastglul00_i=0;
72 if insulin_i>=30 thēn fastins30_i=1;
73 else if $\overline{0}$ <insulin_i<30 then fastins30_i=0;
74

NOTE: There were 6358 observations read from the data set WORK.G6_ST3.
NOTE: There were 6288 observations read from the data set WORK.G6_ST2.

April 29, 2011
NOTE: There were 6173 observations read from the data set WORK.G6_LAB.
NOTE: There were 6313 observations read from the data set WORK.G8_ST3.
NOTE: There were 6187 observations read from the data set WORK.G8_LAB.
NOTE: The data set WORK.BASELINE_F has 4603 observations and 51 variables.
NOTE: DATA statement used (Total process time):

$$
\begin{array}{ll}
\text { real time } & 0.04 \text { seconds } \\
\text { cpu time } & 0.04 \text { seconds }
\end{array}
$$

75 data baseline_f; set baseline_f;
76
77
78
79
80
81

## 82

```
****************** Coding of Final Measures (8th grade) *;
```

    if \(30<\) bmipct_f<85 or bmipct_f=1 then bmige85_f=0;
                else if bmipct_f>=85 or bmipct_f=3 then bmíge85_f=1;
            if \(30<\) bmipct_f<95 or bmipct_f=1 then bmige95_f=0;
                else if bmipct_f>=95 or bmipct_f=3 then bmige95_f=1;
    * recode values >upper limit, or $<l o w e r ~ l i m i t, ~ t o ~ a ~ c o n s t a n t ~ e q u a l ~ t o ~$
exactly that
82 ! limit *;
83 * see forms for values of upper and lower limits *;
84 bmizscore_f=bmiz_f; if bmiz_f=3 then bmizscore_f=2.1; * upper limit
recoded *;
85 if bmiz_f=1 then bmizscore_f=-0.5; * lower limit recoded *;
86 waistcm_f=waist_f; if waist_f=3 then waistcm_f=97;* upper limit recoded
*;
87 if waist_f=1 then waistcm_f=57; * lower limit recoded *;
88
89 if glucose_f>=100 then fastglul00_f=1;
90 else if $\overline{0}<g l u c o s e \_f<100$ then fastglul00_f=0;
91 if insulin_f>=30 thēn fastins30_f=1;
92 else if $\overline{0}$ <insulin_f<30 then fástins30_f=0;
93
94
95
diffbmige85=bmige85_f-bmige85_i;
diffbmige95=bmige95_f-bmige95_i;
diffbmizscore=bmizscore_f-bmizscore_i;
diffwaistcm=waistcm_f-wāistcm_i;
diffglu100=fastglu100_f-fastglu100_i;
diffins30=fastins30_f-fastins30_i;
diffglu=glucose_f-glucose_i;
100
101
102 logins_i=log(insulin_i);
103 if insulin_f=0 then $\bar{l}$ ogins_f=log(1.2); *lowest positive insulin value of
insulin_f*;
104 else logins_f=log(insulin_f);
105 diffins=logins_f-logins_i;
106 diffins_untrans=insulin_f-insulin_i;
107
NOTE: Missing values were generated as a result of performing an operation on missing values.

Each place is given by: (Number of times) at (Line): (Column).
19 at 97:24 221 at 98:26 230 at 99:24 221 at 100:20 142 at 102:12
103 at 104:17 229 at 105:19 229 at 106:28
NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The data set WORK.BASELINE_F has 4603 observations and 68 variables.
NOTE: DATA statement used (Total process time):
real time 0.03 seconds
cpu time 0.01 seconds

```
data baseline_f; set baseline_f;
    famhist=st2famhx1;
    if st2famhx1=9 then famhist=.;
    label bmige85_i='BMI ge 85th percentile, 6th gr'
            bmige95_i='BMI ge 95th percentile, 6th gr'
            famhist='1st deg fam hist diab'
            bmige85_f='BMI ge 85th percentile, 8th gr'
            bmige95_f='BMI ge 95th percentile, 8th gr'
            bmizscore_i='BMI z-score, 6th gr'
            waistcm_i='Waist circum in cm, 6th gr'
            fastglu100_i='Fasting Glucose >=100, 6th gr'
            fastins30_i='Fasting Insulin >=30, 6th gr'
            bmizscore_f='BMI z-score, 8th gr'
            glucose_i='Fasting Glucose, 6th gr'
            insulin_i='Fasting Insulin, 6th gr'
            glucose_f='Fasting Glucose, 8th gr'
            insulin_f='Fasting Insulin, 8th gr'
            waistcm_f='Waist circum in cm, 8th gr'
            fastglu\overline{100_f='Fasting Glucose >=100, 8th gr'}
            fastins30_f='Fasting Insulin >=30, 8th gr'
            diffbmige85='bmige85_f-bmige85_i'
            diffbmige95='bmige95_f-bmige95_i'
            diffbmizscore='bmizscore_f-bmizscore_i'
            diffwaistcm='waistcm_f-waistcm_i'
            diffglu100='fastglu100_f-fastglul00_i'
            diffins30='fastins30_f-fastins30_i'
            diffglu='glucose_f-g\\ucose_i'
            logins_i='Natura\overline{l Log Trans}form of insulin, 6th gr'
            logins_f='Natural Log Transform of insulin, 8th gr'
            diffins='logins_f-logins_i'
            diffins_untrans='insulin_f-insulin_i';
    run;
```

NOTE: There were 4603 observations read from the data set WORK.BASELINE_F. NOTE: The data set WORK.BASELINE_F has 4603 observations and 69 variablés. NOTE: DATA statement used (Total process time):

| real time | 0.01 seconds |
| :--- | :--- |
| cpu time | 0.01 seconds |

140
141
proc contents position; title "Variables in HEALTHY DSIC Analysis
Dataset"; run;
NOTE: PROCEDURE CONTENTS used (Total process time):
real time $\quad 0.28$ seconds
cpu time 0.06 seconds
NOTE: The PROCEDURE CONTENTS printed pages 1-5.
142
143

144
145
146
NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE FREQ printed page 6.
NOTE: PROCEDURE FREQ used (Total process time):
real time 0.03 seconds
cpu time 0.01 seconds

147
proc freq; tables st3gend race_eth bmige85_i st2edu famhist; run;
NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE FREQ printed page 7.
NOTE: PROCEDURE FREQ used (Total process time):
real time 0.18 seconds
cpu time 0.03 seconds

148
149
proc sort; by assign;
NOTE: There were 4603 observations read from the data set WORK.BASELINE_F. NOTE: The data set WORK.BASELINE_F has 4603 observations and 69 variables. NOTE: PROCEDURE SORT used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds

150
proc freq; by assign; tables (st3gend race_eth st2edu famhist )* bmige85_i; run;

NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE FREQ printed pages 8-15.
NOTE: PROCEDURE FREQ used (Total process time):

$$
\text { real time } \quad 0.45 \text { seconds }
$$

cpu time 0.03 seconds

151
152
153
154
155
156

```
            ************************************;
Title To replicate selected results in Table 2 ;
    *************************************;
proc freq; by assign; tables bmige85_i bmige95_i fastglul00_i fastins30_i bmige85_f bmige95_f fastglul00_f fastins30_f; run;
```

NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE FREQ printed pages 16-19.
NOTE: PROCEDURE FREQ used (Total process time):
real time 0.10 seconds
cpu time 0.01 seconds

157 proc means maxdec=4 n mean std stderr; class assign; var bmizscore_i bmizscore_f; run;

NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE MEANS printed page 20.
NOTE: PROCEDURE MEANS used (Total process time):
real time 0.03 seconds
cpu time 0.03 seconds

158 proc means maxdec=4 $n$ mean std stderr; class assign; var waistcm_i
waistcm_f; run;
NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE MEANS printed page 21.
NOTE: PROCEDURE MEANS used (Total process time):
real time 0.09 seconds
cpu time 0.01 seconds

159 proc means maxdec=3 $n$ mean std stderr; class assign; var glucose_i
insulin_i glucose_f
159 ! insulin_f; run;
NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The PROCEDURE MEANS printed page 22.
NOTE: PROCEDURE MEANS used (Total process time): real time 0.03 seconds cpu time 0.01 seconds


162
163
164
165
166
167
168
169
170
171

```
    ***************************************;
Title To replicate analyses of selected continuous outcomes in Table 2 ;
    ***************************************;
proc mixed data=baseline_f;
    class studentid schoolíd assign;
    model diffbmizscore=assign bmizscore_i/solution ddfm=bw;
        random int/sub=schoolid;
        lsmeans assign /diff=control('0') cl; run;
```

WARNING: Class levels for studentid are not printed because of excessive size. NOTE: Convergence criteria met.
NOTE: The PROCEDURE MIXED printed pages 24-26.

April 29, 2011
NOTE: PROCEDURE MIXED used (Total process time):
real time 0.84 seconds
cpu time 0.09 seconds

172
173
174
175
176

```
proc mixed data=baseline_f;
    class studentid schoolíd assign;
    model diffwaistcm=assign waistcm_i/solution ddfm=bw;
    random int/sub=schoolid;
    lsmeans assign /diff=control('0') cl; run;
```

WARNING: Class levels for studentid are not printed because of excessive size. NOTE: 19 observations are not included because of missing values.
NOTE: Convergence criteria met.
NOTE: The PROCEDURE MIXED printed pages 27-29.
NOTE: PROCEDURE MIXED used (Total process time):
real time 0.35 seconds
cpu time 0.04 seconds

177
178
179
180
181

```
proc mixed data=baseline_f;
    class studentid schoolíd assign;
    model diffins=assign logins_i/solution ddfm=bw;
    random int/sub=schoolid;
    lsmeans assign /diff=control('0') cl; run;
```

WARNING: Class levels for studentid are not printed because of excessive size. NOTE: 229 observations are not included because of missing values.
NOTE: Convergence criteria met.
NOTE: The PROCEDURE MIXED printed pages 30-32.
NOTE: PROCEDURE MIXED used (Total process time):
real time 0.15 seconds
cpu time 0.07 seconds

182

```
proc mixed data=baseline_f;
    class studentid schooli}d assign
    model diffglu=assign glucose_i/solution ddfm=bw;
    random int/sub=schoolid;
    lsmeans assign /diff=control('0') cl; run;
```

WARNING: Class levels for studentid are not printed because of excessive size. NOTE: 221 observations are not included because of missing values.
NOTE: Convergence criteria met.
NOTE: The PROCEDURE MIXED printed pages 33-35.
NOTE: PROCEDURE MIXED used (Total process time):
real time 0.35 seconds
cpu time 0.07 seconds

187
188
189

[^1]April 29, 2011

NOTE: There were 4603 observations read from the data set WORK.BASELINE_F.
NOTE: The data set WORK.LONGF has 9206 observations and 74 variables.
NOTE: DATA statement used (Total process time):
real time 0.04 seconds
cpu time 0.01 seconds

```
data longf; set longf;
    time=i-1; run;
```

NOTE: There were 9206 observations read from the data set WORK.LONGF.
NOTE: The data set WORK.LONGF has 9206 observations and 75 variables.
NOTE: DATA statement used (Total process time):
real time 0.21 seconds
cpu time 0.06 seconds

207
208
time i
208 ! diffbmige85; run; */
209
210
211
212
213
214
215

```
data longf; set baseline_f;
    array ge85s[2] bmige85_i bmige85_f;
    array ge95s[2] bmige95_i bmige95_f;
    array fgl00s[2] fastglul00_i fastglul00_f;
    array fi30s[2] fastins30_i fastins30_f;
    do i=1 to 2;
        bmige85=ge85s[i];
            bmige95=ge95s[i];
fastglu100=fg100s[i];
fastins30=fi30s[i];
            output;
        end;
        format bmige85 bmige95 fastglu100 fastins30;
run;
```

| real time | 0.70 seconds |
| :--- | :--- |
| cpu time | 0.54 seconds |

```
PROC GLIMMIX DATA=longf ;
    CLASS studentid schoolid time;
    MODEL bmige95(EVENT=LAST)=assign bmige95_i time /
            DDFM=bw DIST=BINARY cl SOLUTION HTYPE=1,3;
        random int /subject=schoolid;
        run;
```

WARNING: Class levels for studentid are not printed because of excessive size.
NOTE: The GLIMMIX procedure is modeling the probability that bmige95='1'.
NOTE: Convergence criterion (PCONV=1.11022E-8) satisfied.
NOTE: Estimated G matrix is not positive definite.
NOTE: The covariance matrix is the zero matrix.
NOTE: The PROCEDURE GLIMMIX printed pages 39-41.
NOTE: PROCEDURE GLIMMIX used (Total process time):
real time 0.87 seconds
cpu time 0.51 seconds

222
223
224
225
226
227

PROC GLIMMIX DATA=longf ;
CLASS studentid schoolid time;
MODEL fastglul00 (EVENT=LAST) =assign fastglul00_i time /
DDFM=bw DIST=BINARY cl SOLUTION HTYPE=1,3;
random int/subject=schoolid;
run;

WARNING: Class levels for studentid are not printed because of excessive size. NOTE: Some observations are not used in the analysis because of: missing response values ( $n=238$ ),
missing fixed effects ( $\mathrm{n}=270$ ), missing random effects ( $\mathrm{n}=270$ ), missing subject
effects ( $n=270$ ).
NOTE: The GLIMMIX procedure is modeling the probability that fastglu100='1'.
NOTE: Convergence criterion (PCONV=1.11022E-8) satisfied.
NOTE: The PROCEDURE GLIMMIX printed pages 42-44.
NOTE: PROCEDURE GLIMMIX used (Total process time):
real time $\quad 0.65$ seconds
cpu time 0.51 seconds

228
229
230
231
232
233

```
PROC GLIMMIX DATA=longf ;
    CLASS studentid schoolid time;
    MODEL fastins30(EVENT=LAST)=assign fastins30_i time /
            DDFM=bw DIST=BINARY cl SOLUTION HTYPE=1,3;
            random int/subject=schoolid;
            run;
```

April 29, 2011
WARNING: Class levels for studentid are not printed because of excessive size. NOTE: Some observations are not used in the analysis because of: missing response values ( $n=246$ ),
missing fixed effects ( $\mathrm{n}=284$ ), missing random effects ( $\mathrm{n}=284$ ), missing subject effects ( $\mathrm{n}=284$ ).
NOTE: The GLIMMIX procedure is modeling the probability that fastins30='1'.
NOTE: Convergence criterion (PCONV=1.11022E-8) satisfied.
NOTE: The PROCEDURE GLIMMIX printed pages 45-47.
NOTE: PROCEDURE GLIMMIX used (Total process time):
real time 0.70 seconds
cpu time 0.54 seconds

234
235
236
237
NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:
real time $\quad 13.00$ seconds
cpu time 4.68 seconds

## Attachment 3

SAS version 9.2 Output for programming code submitted for the replication of results in Tables 1 and 2 of Healthy Study Group, et al., NEJM 363(5)

Variables in HEALTHY DSIC Analysis Dataset
18:50 Friday, April 29, 2011

The CONTENTS Procedure

| Data Set Name | WORK.BASELINE_F | Observations | 4603 |
| :--- | :--- | :--- | :--- |
| Member Type | DATA | Variables | 69 |
| Engine | V9 | Indexes | 0 |
| Created | Friday, April 29, 2011 06:50:40 PM | Observation Length | 560 |
| Last Modified | Friday, April 29, 2011 06:50:40 PM | Deleted Observations | 0 |
| Protection |  | Compressed | NO |
| Data Set Type |  | Sorted | NO |



| bmizscore_i | Num | 8 |
| :--- | :--- | :--- |
| chol | Num | 8 |
| dbp | Num | 8 |
| diffbmige85 | Num | 8 |
| diffbmige95 | Num | 8 |
| diffbmizscore | Num | 8 |
| diffglu | Num | 8 |
| diffglul00 | Num | 8 |
| diffins | Num | 8 |
| diffins30 | Num | 8 |

BMI z-score, 6th gr
Cholesterol
Diastolic Blood Pressure
bmige85_f-bmige85_i
bmige95_f-bmige95_i
bmizscore f-bmizscore i
glucose_f-glucose_i
fastglu100_f-fastglu100_i
logins_f-lōgins_i
fastins30_f-fastins30_i

The CONTENTS Procedure

Alphabetic List of Variables and Attributes

| Variable | Type | Len |
| :---: | :---: | :---: |
| diffins_untrans | Num | 8 |
| diffwaistcm | Num | 8 |
| famhist | Num | 8 |
| fastglu100_f | Num | 8 |
| fastglu100_i | Num | 8 |
| fastins30_f | Num | 8 |
| fastins30_i | Num | 8 |
| glucose f | Num | 8 |
| glucose_i | Num | 8 |
| grade | Num | 8 |
| hbalc | Num | 8 |
| hdl_chl | Num | 8 |
| height | Num | 8 |
| insulin_f | Num | 8 |
| insulin_i | Num | 8 |
| ldl_chl | Num | 8 |
| logins_f | Num | 8 |
| logins_i | Num | 8 |
| race_eth | Num | 8 |
| s.bp | Num | 8 |
| schoolid | Num | 8 |
| st2comp | Num | 8 |
| st2edu | Num | 8 |
| st2famhx1 | Num | 8 |
| st2famhx2 | Num | 8 |
| st3cuff | Num | 8 |
| st3fast | Num | 8 |
| st3gend | Num | 8 |
| st3infall6 | Num | 8 |
| st3infall7 | Num | 8 |
| st3infall8 | Num | 8 |
| st3inspr6 | Num | 8 |
| st3inspr7 | Num | 8 |
| st3inspr8 | Num | 8 |
| st3reas 1 | Num | 8 |
| st3reas2 | Num | 8 |


| Format | Label |
| :---: | :---: |
|  | insulin_f-insulin_i |
|  | waistcm_f-waistcm_i |
|  | 1st deg fam hist diab |
|  | Fasting Glucose >=100, 8th gr |
|  | Fasting Glucose >=100, 6th gr |
|  | Fasting Insulin $>=30$, 8th gr |
|  | Fasting Insulin >=30, 6th gr |
|  | Fasting Glucose, 8th gr |
|  | Fasting Glucose, 6th gr |
| BEST12. | Grade |
|  | HbA1c |
|  | HDL Cholesterol |
| BEST12. | Height |
|  | Fasting Insulin, 8th gr |
|  | Fasting Insulin, 6th gr |
|  | LDL Cholesterol |
|  | Natural Log Transform of insulin, 8th gr |
|  | Natural Log Transform of insulin, 6th gr |
| BEST12. | Race/Ethnicity |
| BEST12. | Systolic Blood Pressure |
| BEST12. | School ID |
|  | Information completed by |
|  | Highest household education |
|  | First degree family history of diabetes |
|  | Second degree family history of diabetes |
| BEST12. | Cuff Size |
| BEST12. | Fasting Status |
| BEST12. | Gender |
|  | In School Fall 6th Grade |
|  | In School Fall 7th Grade |
|  | In School Fall 8th Grade |
|  | In School Spring 6th Grade |
|  | In School Spring 7th Grade |
|  | In School Spring 8th Grade |
|  | Parent Refused On-site |
|  | Child Refused On-site |

insulin_f-insulin_i
Ist deg fam hist diab
Fasting Glucose >=100, 8th gr
asting Glucose >=100, 6th gr

Fasting Insulin >=30, 6th gr
Fasting Glucose, 8th gr
ing Glucose, 6th gr

HDL Cholesterol

Fasting Insulin, 8th gr
Fasting Insulin, 6th gr
DL Cholesterol
Natural
Race/Ethnicity
Systolic Blood Pressure

Information completed by
First degree family history of diabetes
Second degree family history of diabetes
Size

Gender
In School Fall 6th Grade
7h Grade
S Shool Fall 8th Grade
In School Spring 7th Grade
In School Spring 7th Grade

Refused On-site
Child Refused On-site
st 3 reas 3
st 3 reas 4
st $3 r e a s 5$
st 3 succ
st3tries
studentid
trig
vldl_chl
waist_f
waist_i

Num
Num 8
Num 8
Num 8
Num 8
Char 12
Num 8
Num 8
Num 8
Num
could Not Palpate Vein
3 Tries With No Success
Other Reason
Blood Draw Successful
Number of Blood Draw Tries
Student ID
Triglycerides
VLDL Cholesterol
BEST12. Waist
BEST12. Waist


The CONTENTS Procedure

## Variables in Creation Order

Variable
schoolid
studentid
grade
st3fast
st3gend
race_eth
st3infall6
st3inspr6
st3infall7
st3inspr7
st3infall8
st3inspr8
st3tries
st3succ
st3reas1
st3reas2
st3reas3
st3reas4
st3reas5
height
weight
bmipct_i
bmiz_i
waist i
st3cuff
sbp
dbp
assign
st2comp
st2famhx1
st2famhx2
st2edu
glucose_i
insulin_i
hbalc
chol

| Type | Len | Format | Label |
| :---: | :---: | :---: | :---: |
| Num | 8 | BEST12. | School ID |
| Char | 12 |  | Student ID |
| Num | 8 | BEST12. | Grade |
| Num | 8 | BEST12. | Fasting Status |
| Num | 8 | BEST12. | Gender |
| Num | 8 | BEST12. | Race/Ethnicity |
| Num | 8 |  | In School Fall 6th Grade |
| Num | 8 |  | In School Spring 6th Grade |
| Num | 8 |  | In School Fall 7th Grade |
| Num | 8 |  | In School Spring 7th Grade |
| Num | 8 |  | In School Fall 8th Grade |
| Num | 8 |  | In School Spring 8th Grade |
| Num | 8 | BEST12. | Number of Blood Draw Tries |
| Num | 8 | BEST12. | Blood Draw Successful |
| Num | 8 |  | Parent Refused On-site |
| Num | 8 |  | Child Refused On-site |
| Num | 8 |  | Could Not Palpate Vein |
| Num | 8 |  | 3 Tries With No Success |
| Num | 8 |  | Other Reason |
| Num | 8 | BEST12. | Height |
| Num | 8 | BEST12. | Weight |
| Num | 8 | BEST12. | BMI Percentile |
| Num | 8 |  | BMI Z-score |
| Num | 8 | BEST12. | Waist |
| Num | 8 | BEST12. | Cuff Size |
| Num | 8 | BEST12. | Systolic Blood Pressure |
| Num | 8 | BEST12. | Diastolic Blood Pressure |
| Num | 8 | BEST12. | Randomization Assignment |
| Num | 8 |  | Information completed by |
| Num | 8 |  | First degree family history of diabetes |
| Num | 8 |  | Second degree family history of diabetes |
| Num | 8 |  | Highest household education |
| Num | 8 |  | Fasting Glucose, 6th gr |
| Num | 8 |  | Fasting Insulin, 6th gr |
| Num | 8 |  | HbA1c |
| Num | 8 |  | Cholesterol |


| 37 | hdl_chl | Num | 8 |
| :--- | :--- | :--- | :--- |
| 38 | ldl_chl | Num | 8 |
| 39 | vldl_chl | Num | 8 |
| 40 | trig | Num | 8 |
| 41 | bmipct_f | Num | 8 |
| 42 | bmiz_f | Num | 8 |
| 43 | waist_f | Num | 8 |
| 44 | glucose_f | Num | 8 |
| 45 | insulin_f | Num | 8 |
| 46 | bmige85 i | Num | 8 |

```
    HDL Cholesterol
    LDL Cholesterol
    VLDL Cholesterol
    Triglycerides
    BMI Percentile
    BMI Z-score
    Waist
    Fasting Glucose, 8th gr
    Fasting Insulin, 8th gr
BMIGE85F. BMI ge 85th percentile, 6th gr
```

Variables in HEALTHY DSIC Analysis Dataset

The CONTENTS Procedure

## Variables in Creation Order

| 47 | bmige95_i | Num | 8 |
| :--- | :--- | :--- | :--- |
| 48 | bmizscore_i | Num | 8 |
| 49 | waistcm_i | Num | 8 |
| 50 | fastgluló_i | Num | 8 |
| 51 | fastins30_i | Num | 8 |
| 52 | bmige85_f | Num | 8 |
| 53 | bmige95_f | Num | 8 |
| 54 | bmizscore_f | Num | 8 |
| 55 | waistcm_f | Num | 8 |
| 56 | fastgluloo_f | Num | 8 |
| 57 | fastins30_f | Num | 8 |
| 58 | diffbmige85 | Num | 8 |
| 59 | diffbmige95 | Num | 8 |
| 60 | diffbmizscore | Num | 8 |
| 61 | diffwaistcm | Num | 8 |
| 62 | diffglul00 | Num | 8 |
| 63 | diffins30 | Num | 8 |
| 64 | diffglu | Num | 8 |
| 65 | logins_i | Num | 8 |
| 66 | logins_f | Num | 8 |
| 67 | diffins | Num | 8 |
| 68 | diffins_untrans | Num | 8 |
| 69 | famhist | Num | 8 |

```
Format Label
BMIGE95F. BMI ge 95th percentile, 6th gr
    BMI z-score, 6th gr
    Waist circum in cm, 6th gr
    Fasting Glucose >=100, 6th gr
    Fasting Insulin >=30, 6th gr
    BMI ge 85th percentile, 8th gr
    BMI ge 95th percentile, 8th gr
    BMI z-score, 8th gr
    Waist circum in cm, 8th gr
    Fasting Glucose >=100, 8th gr
    Fasting Insulin >=30, 8th gr
bmige85_f-bmige85_i
bmige95-f-bmige95-i
bmizscore f-bmizscore i
waistcm_f-waistcm_i
fastglu100 f-fastglu100 i
fastins30_\overline{f-fastins30_i}
glucose f-glucose i
Natural Log Transform of insulin, 6th gr
Natural Log Transform of insulin, 8th gr
logins_f-logins_i
insulin f-insulin i
1st deg}\mp@subsup{}{}{-}\mathrm{ fam hist }\overline{\textrm{d}
Label
BMI ge 95th percentile, 6th gr
BMI z-score, 6th gr
aist circum in cm, 6th gr
Fasting Glucose \(>=100\), 6th gr
BMI ge 85th percentile, 8th gr
BMI ge 95th percentile, 8th gr
BMI z-score, 8th gr
Waist circum in cm, 8th gr
Fasting Glucose \(>=100,8\) th \(g\)
Fasting Insulin \(>=30\), 8th \(g r\)
bmige85_f-bmige85_i
bmige95_f-bmige95i
bmizscore_f-bmizscore_i
waistcm_f-waistcm_i
Eastglu100 f-fastglu100 i
fastins30_f-fastins30_i
glucose I-glucose 1
Natural Log Transform of insulin, 6th gr
atural Log Transform of insulin, 8th gr
insulī f-insulin
1st deg \({ }^{-}\)fam hist \(\overline{\mathrm{d}}\) iab
```

To replicate results in Table 1

The FREQ Procedure

Table of assign by bmige85_i
assign (Randomization Assignment)
bmige85 i(BMI ge 85th percentile, 6th gr)


The FREQ Procedure
Gender

| st3gend | Frequency | Percent | Cumulative Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 2175 | 47.25 | 2175 | 47.25 |
| 2 | 2428 | 52.75 | 4603 | 100.00 |

Race/Ethnicity

| race_eth | Frequency | Percent | Cumulative <br> Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| -------------------------------------------------------- |  |  |  |  |
| 1 | 2493 | 54.16 | 2493 | 54.16 |
| 2 | 830 | 18.03 | 3323 | 72.19 |
| 3 | 888 | 19.29 | 4211 | 91.48 |
| 4 | 392 | 8.52 | 4603 | 100.00 |

BMI ge 85th percentile, 6th gr

| bmige85_i | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| :---: | :---: | :---: | :---: | :---: |
| BMI lt 85 | 2311 | 50.21 | 2311 | 50.21 |
| BMI ge 85 | 2292 | 49.79 | 4603 | 100.00 |


| st2edu | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 2310 | 51.67 | 2310 | 51.67 |
| 2 | 2161 | 48.33 | 4471 | 100.00 |

1st deg fam hist diab

| famhist | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 597 | 17.61 | 597 | 17.61 |
| 2 | 2794 | 82.39 | 3391 | 100.00 |

## Randomization Assignment=0

The FREQ Procedure
Table of st3gend by bmige85_i
st 3 gend (Gender)
omige85_i(BMI ge 85th percentile, 6th gr)

Frequency
Percent |
Row Pct |
Col Pct |BMI lt 8|BMI ge 8| Total
$|5 \quad| 5 \quad \mid$

1 | 512 | 570 | 1082
| 22.30 | 24.83 | 47.13
| 47.32 | 52.68 |
| 43.99 | 50.35 |
2 | 652 | 562 |
1214
$|28.40| 24.48 \mid 52.87$
| 53.71 | 46.29 |
| 56.01 | 49.65 |

| ---------+----------------+ |  |  |  |
| :--- | ---: | ---: | ---: |
| Total | 1164 | 1132 | 2296 |

## Randomization Assignment=0

The FREQ Procedure
Table of race_eth by bmige85_i
race_eth (Race/Ethnicity)
bmige85_i(BMI ge 85th percentile, 6th gr)


Randomization Assignment=0
The FREQ Procedure
Table of st2edu by bmige85_i
st2edu(Highest household education)
bmige85 i(BMI ge 85th percentile, 6th gr)

Frequency
Percent |
Row Pct |
Col Pct ${ }^{\mid \mathrm{BMI}}$ lt $8 \mid \mathrm{BMI}$ ge 8| Total


Frequency Missing $=70$

## Randomization Assignment=0

The FREQ Procedure
Table of famhist by bmige85_i
famhist(1st deg fam hist diab)
bmige85_i(BMI ge 85th percentile, 6th gr)

Frequency
Percent |
Row Pct |
Col Pct ${ }^{\mid B M I}$ lt $8 \mid \mathrm{BMI}$ ge 8| Total

308
$\mid 35.06$ | 64.94 |
| 12.40 | 24.18 |
2 - 627
| 44.94 | 36.93 | 81.86
| 54.89 | 45.11 |
| 87.60 | 75.82 |

| ---------+--------+--------+ |  |  |  |
| :--- | ---: | ---: | ---: | ---: |
| Total | 871 | 827 | 1698 |

Frequency Missing $=598$

## Randomization Assignment=1

The FREQ Procedure
Table of st3gend by bmige85_i
st 3 gend (Gender)
omige85_i(BMI ge 85th percentile, 6th gr)

Frequency
Percent |
Row Pct |
Col Pct ${ }_{\mid}^{\mid B M}$ lt $8 \mid \mathrm{BMI}$ ge 8| Total


|  | 1 | 507 | 586 | 1093 |
| :---: | :---: | :---: | :---: | :---: |
|  |  | 21.98 | 25.40 | 47.38 |
|  |  | 46.39 | 53.61 |  |
|  |  | 44.20 | 50.52 |  |
|  | 2 | 640 | 574 | 1214 |
|  |  | 27.74 | 24.88 | 52.62 |
|  |  | 52.72 | 47.28 |  |
|  |  | 55.80 | 49.48 |  |
| Total |  | 1147 | 1160 | 2307 |
|  |  | 49.72 | 50.28 | 100.00 |

## Randomization Assignment=1

The FREQ Procedure
Table of race_eth by bmige85_i
race_eth(Race/Ethnicity)
bmige85_i(BMI ge 85th percentile, 6th gr)


To replicate results in Table 1

## Randomization Assignment=1

The FREQ Procedure

Table of st2edu by bmige85_i
st2edu(Highest household education)
bmige85 i(BMI ge 85th percentile, 6th gr)

Frequency
Percent |
Row Pct |
Col Pct ${ }_{\mid}^{\mid B M}$ lt $8 \mid \mathrm{BMI}$ ge 8| Total
$\mid 5151$

1 | 552 | 611 | 1163
| 24.59 | 27.22 | 51.80
| 47.46 | 52.54 |
| 49.64 | 53.93 |
2 | 560 | 522 |
1082
| 24.94 | 23.25 | 48.20
| 51.76 | 48.24 |
| 50.36 | 46.07 |
$\begin{array}{lccr}\text {----------+---------+--------+ } & \\ \text { Total } & 1112 & 1133 & 2245\end{array}$

Frequency Missing $=62$

## Randomization Assignment=1

The FREQ Procedure
Table of famhist by bmige85_i
famhist(1st deg fam hist diab)
bmige85_i(BMI ge 85th percentile, 6th gr)

Frequency
Percent |
Row Pct |
Col Pct ${ }_{\mid}^{\mid B M}$ lt $8 \mid \mathrm{BMI}$ ge 8| Total

| 1 | 106 | 183 |
| :---: | :---: | :---: |

289
| 6.26 | 10.81 | 17.07
| 36.68 | 63.32 |
| 12.96 | 20.91 |
2 | 712 | 692 |
1404
| 42.06 | 40.87 | 82.93
| 50.71 | 49.29 |
| 87.04 | 79.09 |

| ---------+--------------- |  |  |  |
| :--- | ---: | ---: | ---: |
| Total | 818 | 875 | 1693 |
|  | 48.32 | 51.68 | 100.00 |

Frequency Missing $=614$

Randomization Assignment=0
The FREQ Procedure
BMI ge 85th percentile, 6th gr

| bmige85_i | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| :---: | :---: | :---: | :---: | :---: |
| BMI lt 85 | 1164 | 50.70 | 1164 | 50.70 |
| BMI ge 85 | 1132 | 49.30 | 2296 | 100.00 |

BMI ge 95th percentile, 6th gr

| bmige95_i | Frequency | Percent | Cumulative <br> Frequency | Cumulative <br> Percent |
| :--- | :---: | :---: | :---: | :---: |
| ------------------------------------------------------- |  |  |  |  |
| BMI lt 95 | 1597 | 69.56 | 1597 | 69.56 |
| BMI ge 95 | 699 | 30.44 | 2296 | 100.00 |

Fasting Glucose >=100, 6th gr

| fastglu100_i | Frequency | Percent | Cumulative <br> Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 1870 | 83.74 | 1870 | 83.74 |
| 1 | 363 | 16.26 | 2233 | 100.00 |

Fasting Insulin $>=30$, 6th gr

|  |  | Cumulative | Cumulative |  |
| :---: | :---: | :---: | :---: | :---: |
| fastins30_i | Frequency | Percent | Frequency | Percent |

BMI ge 85th percentile, 8th gr

| bmige85_f | Frequency | Percent | Cumulative <br> Frequency | Cumulative |
| :---: | :---: | :---: | :---: | :---: |
| Percent |  |  |  |  |

To replicate selected results in Table 2 18:50 Friday, April 29, 2011
Randomization Assignment=0
The FREQ Procedure
BMI ge 95th percentile, 8th gr

| bmige95_f | Frequency | Percent | Cumulative <br> Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 1686 | 73.43 | 1686 | 73.43 |
| 1 | 610 | 26.57 | 2296 | 100.00 |

Fasting Glucose >=100, 8th gr

|  |  | Frequency | Percent | Cumulative <br> Frequency |
| :---: | :---: | :---: | :---: | :---: |
| fastglul00_f | Fremulative | Percent |  |  |

Fasting Insulin $>=30$, 8 th $g r$

|  |  | Cumulative | Cumulative |  |
| :---: | :---: | :---: | :---: | :---: |
| fastins30_f | Frequency | Percent | Frequency | Percent |

Randomization Assignment=1
The FREQ Procedure
BMI ge 85th percentile, 6th gr

| bmige85_i | Frequency | Percent | Cumulative <br> Frequency | Cumulative <br> Percent |
| :--- | :---: | :---: | :---: | :---: |
| ---------------------------------------------------------------- |  |  |  |  |
| BMI lt 85 | 1147 | 49.72 | 1147 | 49.72 |
| BMI ge 85 | 1160 | 50.28 | 2307 | 100.00 |

BMI ge 95th percentile, 6th gr

| bmige95_i | Frequency | Percent | Cumulative <br> Frequency | Cumulative <br> Percent |
| :--- | :---: | :---: | :---: | :---: |
| ------------------------------------------------------ |  |  |  |  |
| BMI lt 95 | 1613 | 69.92 | 1613 | 69.92 |
| BMI ge 95 | 694 | 30.08 | 2307 | 100.00 |

Fasting Glucose >=100, 6th gr

| fastglu100_i | Frequency | Percent | Cumulative Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 1882 | 84.21 | 1882 | 84.21 |
| 1 | 353 | 15.79 | 2235 | 100.00 |

Fasting Insulin $>=30$, 6th gr

|  |  | Cumulative | Cumulative |  |
| :---: | :---: | :---: | :---: | :---: |
| fastins30_i | Frequency | Percent | Frequency | Percent |

BMI ge 85th percentile, 8th gr

| bmige85_f | Frequency | Percent | Cumulative Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 1250 | 54.18 | 1250 | 54.18 |
| 1 | 1057 | 45.82 | 2307 | 100.00 |

To replicate selected results in Table 2 18:50 Friday, April 29, 2011
Randomization Assignment=1
The FREQ Procedure
BMI ge 95th percentile, 8th gr

| bmige95_f | Frequency | Percent | Cumulative Frequency | Cumulative <br> Percent |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 1740 | 75.42 | 1740 | 75.42 |
| 1 | 567 | 24.58 | 2307 | 100.00 |

Fasting Glucose >=100, 8th gr

|  |  | Frequency | Percent | Cumulative <br> Frequency |
| :---: | :---: | :---: | :---: | :---: |
| fastglul00_f | Frumulative |  |  |  |
| Percent |  |  |  |  |

Fasting Insulin $>=30$, 8th gr

|  |  | Cumulative | Cumulative |  |
| :---: | :---: | :---: | :---: | :---: |
| fastins30_f | Frequency | Percent | Frequency | Percent |

To replicate selected results in Table 2 18:50 Friday, April 29, 2011
The MEANS Procedure

| Randomization Assignment | $\begin{array}{r} \mathrm{N} \\ \text { Obs } \end{array}$ | Variable | Label |  | N | Mean | Std Dev |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 2296 | bmizscore i | BMI z-score, | 6 th gr | 2296 | 0.9161 | 0.9221 |
|  |  | bmizscore_f | BMI z-score, | 8th gr | 2296 | 0.8853 | 0.8810 |


| Randomization Assignment | $\begin{array}{r} \mathrm{N} \\ \text { Obs } \end{array}$ | Variable | Label |  | Std Error |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 2296 | bmizscore i | BMI z-score, | 6 th gr | 0.0192 |
|  |  | bmizscore_f | BMI z-score, | 8th gr | 0.0184 |
| 1 | 2307 | bmizscore_i | BMI z-score, | 6 th gr | 0.0190 |
|  |  | bmizscore_f | BMI z-score, | 8th gr | 0.0182 |

To replicate selected results in Table 2 18:50 Friday, April 29, 2011
The MEANS Procedure



The MEANS Procedure


| Randomization Assignment | $\begin{array}{r} \mathrm{N} \\ \text { Obs } \end{array}$ | Variable | Label | Std Error |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 2296 | glucose_i | Fasting Glucose, 6th gr | 0.143 |
|  |  | insulin_i | Fasting Insulin, 6th gr | 0.261 |
|  |  | glucose_f | Fasting Glucose, 8th gr | 0.167 |
|  |  | insulin_f | Fasting Insulin, 8th gr | 0.281 |
| 1 | 2307 | glucose_i | Fasting Glucose, 6th gr | 0.139 |
|  |  | insulin_i | Fasting Insulin, 6th gr | 0.225 |
|  |  | glucose_f | Fasting Glucose, 8th gr | 0.181 |
|  |  | insulin_f | Fasting Insulin, 8th gr | 0.321 |

The MEANS Procedure

| Randomization Assignment | $\begin{array}{r} \mathrm{N} \\ \text { Obs } \end{array}$ | Variable | Label | N | Mean | Std Dev |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 2296 | diffbmige85 | bmige85_f-bmige85_i | 2296 | -0.041 | 0.337 |
|  |  | diffbmige95 | bmige95-f-bmige95_i | 2296 | -0.039 | 0.320 |
|  |  | diffbmizscore | bmizscore_f-bmizscoore_i | 2296 | -0.031 | 0.390 |
|  |  | diffwaistcm | waistcm_f-waistcm_i | 2287 | 4.030 | 6.079 |
|  |  | diffglul00 | fastglu100_f-fastglu100_i | 2186 | 0.069 | 0.473 |
|  |  | diffins30 | fastins30_f-fastins30_i | 2186 | 0.039 | 0.326 |
|  |  | diffglu | glucose_f-glucose_i | 2186 | 0.699 | 7.839 |
|  |  | diffins_untrans | insulin_f-insulin_i | 2186 | 4.029 | 12.426 |
| 1 | 2307 | diffbmige85 | bmige85_f-bmige85_i | 2307 | -0.045 | 0.345 |
|  |  | diffbmige95 | bmige95_f-bmige95_i | 2307 | -0.055 | 0.323 |
|  |  | diffbmizscore | bmizscore_f-bmizscoore_i | 2307 | -0.062 | 0.405 |
|  |  | diffwaistcm | waistcm_f-waistcm_i | 2297 | 3.493 | 5.971 |
|  |  | diffglul00 | fastglu100_f-fastglu100_i | 2196 | 0.049 | 0.469 |
|  |  | diffins30 | fastins30_f-fastins30_i | 2187 | 0.038 | 0.311 |
|  |  | diffglu | glucose_f-glucose_i | 2196 | 0.023 | 8.592 |
|  |  | diffins_untrans | insulin_f-insulin_i | 2188 | 3.814 | 14.155 |


| Randomization Assignment | $\begin{array}{r} \mathrm{N} \\ \text { Obs } \end{array}$ | Variable | Label | Std Error |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 2296 | diffbmige85 | bmige85_f-bmige85_i | 0.007 |
|  |  | diffbmige95 | bmige95-f-bmige95-i | 0.007 |
|  |  | diffbmizscore | bmizscore_f-bmizscore_i | 0.008 |
|  |  | diffwaistcm | waistcm_f-waistcm_i | 0.127 |
|  |  | diffglu100 | fastglu100_f-fastglul00_i | 0.010 |
|  |  | diffins30 | fastins30_f-fastins30_i | 0.007 |
|  |  | diffglu | glucose_f-glucose_i | 0.168 |
|  |  | diffins_untrans | insulin_f-insulin_i | 0.266 |
| 1 | 2307 | diffbmige85 | bmige85_f-bmige85_i | 0.007 |
|  |  | diffbmige95 | bmige95_f-bmige95_i | 0.007 |
|  |  | diffbmizscore | bmizscore_f-bmizscore_i | 0.008 |
|  |  | diffwaistcm | waistcm_f-waistcm_i | 0.125 |
|  |  | diffglu100 | fastglul00_f-fastglul00_i | 0.010 |
|  |  | diffins30 | fastins30_f-fastins30_i | 0.007 |
|  |  | diffglu | glucose_f-glucose_i | 0.183 |

To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Model Information


## Dimensions

| Covariance Parameters | 2 |
| :--- | ---: |
| Columns in X | 4 |
| Columns in Z Per Subject | 1 |
| Subjects | 42 |
| Max Obs Per Subject | 179 |

## Number of Observations

| Number of Observations Read | 4603 |
| :--- | ---: |
| Number of Observations Used | 4603 |
| Number of Observations Not Used | 0 |

## Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
| ---: | ---: | ---: | ---: |
| 0 | 1 | 4115.30441025 |  |
| 1 | 2 | 4105.88182285 | 0.00000000 |

To replicate analyses of selected continuous outcomes in Table 2

The Mixed Procedure

Convergence criteria met.

| Covariance | Parameter | Estimates |
| :--- | ---: | ---: |
| Cov Parm | Subject | Estimate |
| Intercept | schoolid | 0.001071 |
| Residual |  | 0.1415 |

## Fit statistics

| -2 Res Log Likelihood | 4105.9 |
| :--- | :--- |
| AIC (smaller is better) | 4109.9 |
| AICC (smaller is better) | 4109.9 |
| BIC (smaller is better) | 4113.4 |

Solution for Fixed Effects


|  | Randomization |  | Standard |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Effect | Assignment | Estimate | Error | DF | t Value | $\operatorname{Pr}>\|t\|$ | Alpha | Lower | Upper |
| assign | 0 | -0.03103 | 0.01073 | 40 | -2.89 | 0.0062 | 0.05 | -0.05272 | -0.00934 |
| assign | 1 | -0.05855 | 0.01070 | 40 | -5.47 | <.0001 | 0.05 | -0.08017 | -0.03693 |

To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Differences of Least Squares Means

| Effect | Randomization Assignment | Randomization Assignment | Estimate | Standard Error | DF | t Value | > ${ }^{\text {t }}$ | a |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| assign | 1 | 0 | -0.02752 | 0.01515 | 40 | -1.82 | 0.0769 | 0.05 |
| Differences of Least Squares Means |  |  |  |  |  |  |  |  |
|  | Effect | Randomization Assignment | Randomization |  | Lower | Upper |  |  |
|  | assign | 1 |  | 0 | 05815 | 0.0031 |  |  |

To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Model Information


## Dimensions

| Covariance Parameters | 2 |
| :--- | ---: |
| Columns in X | 4 |
| Columns in Z Per Subject | 1 |
| Subjects | 42 |
| Max Obs Per Subject | 179 |

## Number of Observations

| Number of Observations Read | 4603 |
| :--- | :--- |
| Number of Observations Used | 4584 |
| Number of Observations Not Used | 19 |

Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
| ---: | ---: | ---: | ---: |
| 0 | 1 | 28805.72900680 |  |
| 1 | 3 | 28743.57025603 | 0.00000029 |
| 2 | 1 | 28743.56721191 | 0.00000000 |

To replicate analyses of selected continuous outcomes in Table 2

The Mixed Procedure

Convergence criteria met.

| Covariance | Parameter | Estimates |
| :--- | ---: | ---: |
| Cov Parm | Subject | Estimate |
|  |  |  |
| Intercept | schoolid | 0.7780 |
| Residual |  | 30.5365 |

## Fit Statistics

| -2 Res Log Likelihood | 28743.6 |
| :--- | :--- |
| AIC (smaller is better) | 28747.6 |
| AICC (smaller is better) | 28747.6 |
| BIC (smaller is better) | 28751.0 |

Solution for Fixed Effects



To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Differences of Least Squares Means


To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Model Information


## Dimensions

| Covariance Parameters | 2 |
| :--- | ---: |
| Columns in X | 4 |
| Columns in Z Per Subject | 1 |
| Subjects | 42 |
| Max Obs Per Subject | 179 |

## Number of Observations

| Number of Observations Read | 4603 |
| :--- | ---: |
| Number of Observations Used | 4374 |
| Number of Observations Not Used | 229 |

## Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
| ---: | ---: | ---: | ---: |
| 0 | 1 | 6536.14089549 |  |
| 1 | 2 | 6513.28241995 | 0.00000000 |

To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Convergence criteria met.

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
| :--- | ---: | ---: |
| Intercept | schoolid | 0.003592 |
| Residual |  | 0.2564 |

Fit Statistics

| -2 Res Log Likelihood | 6513.3 |
| :--- | :--- |
| AIC (smaller is better) | 6517.3 |
| AICC (smaller is better) | 6517.3 |
| BIC (smaller is better) | 6520.8 |

Solution for Fixed Effects

| Effect | Randomization Assignment | Estimate | Standard Error | DF | t Value | Pr > \|t| |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept |  | 1.4628 | 0.03172 | 40 | 46.12 | $<.0001$ |
| assign | 0 | 0.05317 | 0.02429 | 40 | 2.19 | 0.0345 |
| assign | 1 | 0 | . | - | . | . |
| logins_i |  | -0.5086 | 0.01139 | 4331 | -44.65 | <. 0001 |


|  | Type 3 Tests of Fixed Effects |  |  |  |
| :--- | ---: | ---: | ---: | ---: |
|  | Num | Den |  |  |
| Effect | DF | DF | F Value | Pr $>$ F |
| assign | 1 | 40 | 4.79 | 0.0345 |
| logins_i | 1 | 4331 | 1993.23 | $<.0001$ |
|  |  |  |  |  |
|  | Least Squares Means |  |  |  |


|  | Randomization |  | Standard |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Effect | Assignment | Estimate | Error | DF | t Value | $\operatorname{Pr}>\|t\|$ | Alpha | Lower | Upper |
| assign | 0 | 0.3270 | 0.01720 | 40 | 19.01 | $<.0001$ | 0.05 | 0.2922 | 0.3618 |
| assign | 1 | 0.2738 | 0.01716 | 40 | 15.96 | <.0001 | 0.05 | 0.2392 | 0.3085 |

To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Differences of Least Squares Means


To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Model Information


## Dimensions

| Covariance Parameters | 2 |
| :--- | ---: |
| Columns in X | 4 |
| Columns in Z Per Subject | 1 |
| Subjects | 42 |
| Max Obs Per Subject | 179 |

## Number of Observations

| Number of Observations Read | 4603 |
| :--- | ---: |
| Number of Observations Used | 4382 |
| Number of Observations Not Used | 221 |

Iteration History

| Iteration | Evaluations | -2 Res Log Like | Criterion |
| ---: | ---: | ---: | ---: |
| 0 | 1 | 30158.26845337 |  |
| 1 | 2 | 30011.13350166 | 0.00000005 |
| 2 | 1 | 30011.13291181 | 0.00000000 |

To replicate analyses of selected continuous outcomes in Table 2

The Mixed Procedure

Convergence criteria met.

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
| :--- | :--- | ---: |
| Intercept | schoolid | 2.8565 |
| Residual |  | 54.1901 |

Fit Statistics

| -2 Res Log Likelihood | 30011.1 |
| :--- | :--- |
| AIC (smaller is better) | 30015.1 |
| AICC (smaller is better) | 30015.1 |
| BIC (smaller is better) | 30018.6 |

Solution for Fixed Effects

| Effect | Randomization Assignment | Estimate | Standard Error | DF | t Value | Pr > \|t| |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept |  | 46.2798 | 1.6429 | 40 | 28.17 | $<.0001$ |
| assign | 0 | 0.5676 | 0.5709 | 40 | 0.99 | 0.3261 |
| assign | 1 | 0 |  | - | . |  |
| glucose_i |  | -0.4953 | 0.01708 | 4339 | -29.00 | <.0001 |


|  | Type 3 Tests of Fixed Effects |  |  |  |
| :--- | ---: | ---: | ---: | ---: |
|  | Num | Den |  |  |
| Effect | DF | DF | F Value | Pr $>$ F |
| assign | 1 | 40 | 0.99 | 0.3261 |
| glucose_i | 1 | 4339 | 841.11 | $<.0001$ |
|  |  |  |  |  |
|  | Least Squares Means |  |  |  |


|  | Randomization <br> Assignment | Estimate | Standard <br> Error | DF | $t$ Value | Pr $>\|t\|$ | Alpha | Lower | Upper |
| :--- | :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| assign | 0 | 0.5263 | 0.4040 | 40 | 1.30 | 0.2002 | 0.05 | -0.2903 | 1.3428 |
| assign | 1 | -0.04134 | 0.4033 | 40 | -0.10 | 0.9189 | 0.05 | -0.8565 | 0.7738 |

To replicate analyses of selected continuous outcomes in Table 2
18:50 Friday, April 29, 2011

The Mixed Procedure

Differences of Least Squares Means


To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Model Information

| Data Set | WORK.LONGF |
| :--- | :--- |
| Response Variable | bmige85 |
| Response Distribution | Binary |
| Link Function | Logit |
| Variance Function | Default |
| Variance Matrix Blocked By | Schoolid |
| Estimation Technique | Residual PL |
| Degrees of Freedom Method | Between-Within |

Class Level Information

| Class | Levels | Values |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| studentid | 4603 | not printed |  |  |  |  |  |  |  |  |  |  |  |
| schoolid | 42 | $\begin{array}{lllll}11 & 13 & 15 & 16\end{array}$ | 18 | 21 | 23 | 24 | 25 | 26 | 27 | 33 | 36 | 38 | 39 |
|  |  | 40414548 | 51 | 53 | 55 | 56 | 57 | 60 | 63 | 66 | 70 | 71 |  |
|  |  | $\begin{array}{llll}73 & 74 & 76 & 77\end{array}$ | 79 | 86 | 87 | 88 | 90 | 94 | 96 | 98 |  |  |  |
| time | 2 | 01 |  |  |  |  |  |  |  |  |  |  |  |


| Number of Observations Read | 9206 |
| :--- | :--- |
| Number of Observations Used | 9206 |


| Response Profile |  |  |
| :---: | :--- | ---: |
| Ordered |  | Total |
| Value | bmige85 | Frequency |
| 1 | 0 | 4820 |
| 2 | 1 | 4386 |

The GLIMMIX procedure is modeling the probability that bmige85='1'.

| G-side Cov. Parameters | 1 |
| :--- | ---: |
| Columns in X | 5 |
| Columns in Z per Subject | 1 |
| Subjects (Blocks in V) | 42 |
| Max Obs per Subject | 358 |

To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Optimization Information

| Optimization Technique | Newton-Raphson with Ridging |
| :--- | :--- |
| Parameters in Optimization | 1 |
| Lower Boundaries | 1 |
| Upper Boundaries | 0 |
| Fixed Effects | Profiled |
| Starting From | Data |

## Iteration History



Standard

| Cov Parm | Subject | Estimate | Error |
| :--- | :---: | :---: | :---: |
| Intercept | schoolid | 0 | . |

To replicate analyses of selected dichotomous outcomes in Table 2
18:50 Friday, April 29, 2011
The GLIMMIX Procedure

Solutions for Fixed Effects

| Effect | time | Estimate | Standard Error | DF | t Value | $\operatorname{Pr}>\|t\|$ | Alpha | Lower | Upper |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept |  | -3.7258 | 0.1103 | 40 | -33.79 | $<.0001$ | 0.05 | -3.9486 | -3.5029 |
| assign |  | -0.00873 | 0.08921 | 40 | -0.10 | 0.9225 | 0.05 | -0.1890 | 0.1716 |
| bmige85_i |  | 5.8208 | 0.1004 | 9162 | 57.96 | <. 0001 | 0.05 | 5.6240 | 6.0177 |
| time | 0 | 0.8154 | 0.09491 | 41 | 8.59 | <. 0001 | 0.05 | 0.6237 | 1.0071 |
| time | 1 | 0 |  | . | - | . | . | - |  |

Type I Tests of Fixed Effects


To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Model Information

| Data Set | WORK.LONGF |
| :--- | :--- |
| Response Variable | bmige95 |
| Response Distribution | Binary |
| Link Function | Logit |
| Variance Function | Default |
| Variance Matrix Blocked By | Schoolid |
| Estimation Technique | Residual PL |
| Degrees of Freedom Method | Between-Within |

Class Level Information

| Class | Levels | Values |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| studentid | 4603 | not printed |  |  |  |  |  |  |  |  |  |  |  |
| schoolid | 42 | $\begin{array}{lllll}11 & 13 & 15 & 16\end{array}$ | 18 | 21 | 23 | 24 | 25 | 26 | 27 | 33 | 36 | 38 | 39 |
|  |  | 40414548 | 51 | 53 | 55 | 56 | 57 | 60 | 63 | 66 | 70 | 71 |  |
|  |  | $\begin{array}{llll}73 & 74 & 76 & 77\end{array}$ | 79 | 86 | 87 | 88 | 90 | 94 | 96 | 98 |  |  |  |
| time | 2 | 01 |  |  |  |  |  |  |  |  |  |  |  |


| Number of Observations Read | 9206 |
| :--- | :--- |
| Number of Observations Used | 9206 |

## Response Profile

## Ordered

 Value bmige 95 0| 1 | 0 |
| :--- | :--- |
| 2 | 1 |

Total
Frequency

6636
2570

The GLIMMIX procedure is modeling the probability that bmige95='1'.

| G-side Cov. Parameters | 1 |
| :--- | ---: |
| Columns in X | 5 |
| Columns in Z per Subject | 1 |
| Subjects (Blocks in V) | 42 |
| Max Obs per Subject | 358 |

To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Optimization Information

| Optimization Technique | Newton-Raphson with Ridging |
| :--- | :--- |
| Parameters in Optimization | 1 |
| Lower Boundaries | 1 |
| Upper Boundaries | 0 |
| Fixed Effects | Profiled |
| Starting From | Data |

## Iteration History



Standard

| Cov Parm | Subject | Estimate | Error |
| :--- | :---: | :---: | :---: |
| Intercept | schoolid | 0 | . |

To replicate analyses of selected dichotomous outcomes in Table 2
18:50 Friday, April 29, 2011

The GLIMMIX Procedure

Solutions for Fixed Effects

| Effect | time | Estimate | Standard Error | DF | t Value | $\operatorname{Pr}>\|t\|$ | Alpha | Lower | Upper |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept |  | -4.4052 | 0.1246 | 40 | -35.35 | $<.0001$ | 0.05 | -4.6571 | -4.1533 |
| assign |  | -0.1861 | 0.09672 | 40 | -1.92 | 0.0615 | 0.05 | -0.3815 | 0.009416 |
| bmige95_i |  | 6.0135 | 0.1141 | 9162 | 52.71 | <. 0001 | 0.05 | 5.7898 | 6.2371 |
| time | 0 | 1.0526 | 0.1057 | 41 | 9.96 | <. 0001 | 0.05 | 0.8391 | 1.2662 |
| time | 1 | 0 |  | . | . | . | . |  |  |

Type I Tests of Fixed Effects


To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Model Information

| Data Set | WORK.LONGF |
| :--- | :--- |
| Response Variable | fastgluloo |
| Response Distribution | Binary |
| Link Function | Logit |
| Variance Function | Default |
| Variance Matrix Blocked By | Schoolid |
| Estimation Technique | Residual PL |
| Degrees of Freedom Method | Between-Within |

Class Level Information


The GLIMMIX procedure is modeling the probability that fastglul00='1'.

| G-side Cov. Parameters | 1 |
| :--- | ---: |
| Columns in X | 5 |
| Columns in Z per Subject | 1 |
| Subjects (Blocks in V) | 42 |
| Max Obs per Subject | 356 |

To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Optimization Information

| Optimization Technique | Newton-Raphson with Ridging |
| :--- | :--- |
| Parameters in Optimization | 1 |
| Lower Boundaries | 1 |
| Upper Boundaries | 0 |
| Fixed Effects | Profiled |
| Starting From | Data |

## Iteration History



## Solutions for Fixed Effects

| Effect | time | Estimate | Error | DF | t Value | Pr $>\|t\|$ | Alpha | Lower |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
|  |  |  |  |  |  |  |  |  |  |
| Intercept |  | -2.1500 | 0.1076 | 40 | -19.97 | $<.0001$ | 0.05 | -2.3675 | -1.9324 |
| assign |  | -0.08728 | 0.1435 | 40 | -0.61 | 0.5464 | 0.05 | -0.3772 | 0.2027 |
| fastglul00_i |  | 3.5669 | 0.07888 | 8806 | 45.22 | $<.0001$ | 0.05 | 3.4123 | 3.7216 |
| time | 0 | -0.6367 | 0.07163 | 41 | -8.89 | $<.0001$ | 0.05 | -0.7813 | -0.4920 |

To replicate analyses of selected dichotomous outcomes in Table 2
18:50 Friday, April 29, 2011

The GLIMMIX Procedure

Solutions for Fixed Effects

| Effect | time | Estimate | Error | DF | $t$ | Value | Alpha | Lower | Upper |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| time | 1 | 0 |  |  |  |  |  |  |  |

ype I Tests of Fixed Effects

|  | Num | Den |  |  |
| :--- | ---: | ---: | ---: | ---: |
| Effect | DF | DF | F Value | Pr $>$ F |
| assign | 1 | 40 | 0.51 | 0.4780 |
| fastglul00_i | 1 | 8806 | 2044.89 | $<.0001$ |
| time | 1 | 41 | 79.00 | $<.0001$ |


| Type III Tests of Fixed Effects |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: |
|  | Num | Den |  |  |
| Effect | DF | DF | F Value | Pr $>$ F |
|  |  |  |  |  |
| assign | 1 | 40 | 0.37 | 0.5464 |
| fastglul00_i | 1 | 8806 | 2044.90 | $<.0001$ |
| time | 1 | 41 | 79.00 | $<.0001$ |

```
To replicate analyses of selected dichotomous outcomes in Table 2
                                    18:50 Friday, April 29, 2011
The GITMMIX Procedure
Model Information
\begin{tabular}{ll} 
Data Set & WORK.LONGF \\
Response Variable & fastins30 \\
Response Distribution & Binary \\
Link Function & Logit \\
Variance Function & Default \\
Variance Matrix Blocked By & schoolid \\
Estimation Technique & Residual PL \\
Degrees of Freedom Method & Between-Within
\end{tabular}
Class Level Information
```



```
\begin{tabular}{ll} 
Number of Observations Read & 9206 \\
Number of Observations Used & 8834
\end{tabular}
```


## Response Profile

Ordered Value fastins30 Frequency

10
8063
771

```
The GLIMMIX procedure is modeling the probability that fastins30='1'.
```

| G-side Cov. Parameters | 1 |
| :--- | ---: |
| Columns in X | 5 |
| Columns in Z per Subject | 1 |
| Subjects (Blocks in V) | 42 |
| Max Obs per Subject | 356 |

To replicate analyses of selected dichotomous outcomes in Table 2

The GLIMMIX Procedure

Optimization Information

| Optimization Technique | Newton-Raphson with Ridging |
| :--- | :--- |
| Parameters in Optimization | 1 |
| Lower Boundaries | 1 |
| Upper Boundaries | 0 |
| Fixed Effects | Profiled |
| Starting From | Data |

## Iteration History



| Effect | time | Estimate | Standard Error | DF | t Value | Pr > \|t| | Alpha | Lower | Upper |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Intercept |  | -2.8465 | 0.08919 | 40 | -31.92 | $<.0001$ | 0.05 | -3.0267 | -2.6662 |
| assign |  | -0.07829 | 0.1106 | 40 | -0.71 | 0.4830 | 0.05 | -0.3018 | 0.1452 |
| fastins30_i |  | 4.5599 | 0.1189 | 8790 | 38.35 | $<.0001$ | 0.05 | 4.3269 | 4.7930 |

To replicate analyses of selected dichotomous outcomes in Table 2
18:50 Friday, April 29, 2011

The GLIMMIX Procedure

Solutions for Fixed Effects

| Effect | time | Estimate | Error <br> Er | DF | $t$ Value | Pr $>\|t\|$ | Alpha | Lower | Upper |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| time | 0 | -0.8774 | 0.1066 | 41 | -8.23 | $<.0001$ | 0.05 | -1.0926 | -0.6621 |
| time | 1 | 0 | . | . | . | . | . | . |  |

Type I Tests of Fixed Effects

|  | Num | Den |  |  |
| :--- | ---: | ---: | ---: | ---: |
| Effect | DF | DF | F Value | Pr $>$ F |
| assign | 1 | 40 | 1.19 | 0.2812 |
| fastins30_i | 1 | 8790 | 1471.28 | $<.0001$ |
| time | 1 | 41 | 67.76 | $<.0001$ |

Type III Tests of Fixed Effects

|  | Num | Den |  |  |
| :--- | ---: | ---: | ---: | ---: |
| Effect | DF | DF | F Value | Pr $>$ F |
| assign |  |  |  | 0.50 |
| fastins30_i | 1 | 40 | 0.4830 |  |
| time | 1 | 8790 | 1470.98 | $<.0001$ |
|  | 1 | 41 | 67.76 | $<.0001$ |


[^0]:    HEALTHY Study Group, Foster GD, Linder B, Baranowski T, Cooper DM, Goldberg L, Harrell JS, Kaufman F, Marcus MD, Treviño RP, Hirst K. A school-based intervention for diabetes risk reduction. New England Journal of Medicine 363(5) [2010 Jul 29]: 443-53.

    Epub 2010 Jun 27.

[^1]:    ************************************;
    Title To replicate analyses of selected dichotomous outcomes in Table 2 ;

