Norma Pugh
May 12, 2006

Integrity Check for The Diabetes Prevention Type 1 (DPT-1) Files

As a partial check of the integrity of the DPT-1 datasets archived in the NIDDK data repository, a set of tabulations was performed to verify that published results from the DPT-1 study can be reproduced using the archived datasets. A small number of analyses were performed to duplicate published results for the data reported by the Type 1 Diabetes Study Group [1] in Diabetes Care in May, 2005 . The results of this integrity check are described below. The full text of the Diabetes Care article can be found in Attachment 1, and the SAS code for our tabulations is included in Attachment 2.

In general, sample sizes, demographic data, baseline measurements and corresponding p-values were successfully replicated. Many of the follow-up measurements and selected baseline measurements can not be successfully replicated without additional explanation of how these variables were defined. The DCC has agreed to participate in a meeting to discuss these issues, but has also stated that all published results were independently confirmed.

Data and Structures. In the data structure of the DPT-1 study, there is no single baseline dataset. The following datasets were used for the integrity check of the baseline tables (Table 1. Baseline characteristics of randomly assigned subjects; Table 2. Baseline characteristics of subjects by IAA status):

SUBJECT	Subject Table (contains race, gender, treatment assignment, etc.)
TEST	Test Table (contains all test results for DPT-1)
RELATIVE	Relative Table (contains information on subject's probands)
CODETABLE	Code groups for all coded variables

Sample size was obtained by restricting the SUBJECT dataset to observations with a value of 'Oral Insulin' or 'Oral Placebo' for the *treatment* variable.

Age in years was calculated as the value of the *age_at_randomization* variable (SUBJECT dataset) divided by 365.25.

First-phase insulin response (FPIR) was obtained by adding together the insulin values at one and three minutes (TEST dataset, *ins1* and *ins3* variables, where the *testname* variable has a value of 'AB-IVGTT' or 'CO-IVGTT'). Insulin values were transformed from character to numeric, with values of the form '<4', '<3', etc. recoded to zero. The definition of first-phase insulin response is documented in the Research Design and Methods: Tolerance test procedures section of the Skyler article.

Race and sex variables are found in the SUBJECT dataset (variable names *race* and *gender*).

Relationship to index patient with diabetes is captured in the RELATIVE dataset (variable name *reltype*). Relationships were collapsed into categories of sibling, offspring, parent, or second degree.

Values for ICA antibodies were obtained from the *result* variable of the TEST dataset, where the *testname* variable has a value of 'ICA'. ICA values were transformed from character to numeric.

Values for IAA antibodies were obtained from the *result* variable of the TEST dataset, where the *testname* variable has a value of 'IAA'. IAA values were transformed from character to numeric.

GAD – *unable to find this variable.*

ICA-512 – *unable to find this variable.*

Micro IAA – *unable to find this variable.*

Values for HbA_{1c} were obtained from the *result* variable of the TEST dataset, where the *testname* variable has a value of 'HBA1C'. HbA_{1c} values were transformed from character to numeric.

C-peptide area under curve – need to know what method was used.

Subjects may have multiple baseline observations. To identify a single baseline record for each subject, all observations with a reported age less than or equal to the *age_at_randomization* variable (SUBJECT dataset) were kept (Note: study timepoints are identified by subject's age in days). Datasets were then sorted by ID and age. The last observation (i.e., the observation occurring closest to the day of randomization but not after randomization) was then kept as the baseline observation for a subject. The exception to this rule is HbA_{1c}, for which the first observation for each subject was used as the baseline observation.

Table 2 reports the same results as Table 1, stratified by baseline IAA values. Subjects with IAA values ≥ 80 nU/ml (confirmed on two occasions) were compared to those with IAA values not confirmed ≥ 80 nU/ml.

Table 1: Comparison of Values of Baseline Characteristics of Randomly Assigned Subjects Computed in Integrity Check to Reference Article Values

Variable	Oral Insulin Group			Oral Placebo Group			P-value		
	Skyler et al	Integrity check	Diff	Skyler et al	Integrity check	Diff	Skyler et al	Integrity check	Diff
N	186	186	0	186	186	0			
Median age	11.0 (7 - 14)	11.1 (7 – 14)	0.1	9.5 (7 – 14)	10.0 (7 – 14)	0.5	0.3569	0.3950	0.0381
Average FPIR (μU/ml)	161.6 ± 72.4	161.5 ± 72.4	0.1	158.9 ± 99.2	158.9 ± 99.2	0	0.7672	0.7672	0
Race:							0.2807	0.3905	0.1098
White	164 (88.1)	164 (88.2)	0 (0.1)	163 (87.6)	163 (87.6)	0			
African American	5 (2.6)	5 (2.7)	0 (0.1)	2 (1.0)	2 (1.1)	0 (0.1)			
Hispanic	8 (4.3)	8 (4.3)	0	14 (7.5)	14 (7.5)	0			
Other	9 (4.7)	9 (4.8)	0 (0.1)	7 (3.7)	7 (3.8)	0 (0.1)			
Sex:							0.1381	0.1381	0
Male	119 (63.9)	119 (64.0)	0 (0.1)	105 (56.4)	105 (56.5)	0 (0.1)			
Female	67 (36.0)	67 (36.0)	0	81 (43.5)	81 (43.6)	0 (0.1)			
Relationship to index patient w/ diabetes							0.6552	0.6835	0.0283
Sibling	112 (60.2)	112 (60.2)	0	108 (58.0)	108 (58.1)	0 (0.1)			
Offspring	49 (26.3)	49 (26.3)	0	53 (28.4)	54 (29.0)	1 (0.6)			
Parent	11 (5.9)	11 (5.9)	0	7 (3.7)	7 (3.8)	0 (0.1)			
Second degree	14 (7.5)	14 (7.5)	0	18 (9.6)	17 (9.1)	1 (0.5)			
Antibody levels:									
Median ICAs (JDF units)	80 (403 – 20)	80 (40 – 320)	0*	80 (40 – 160)	80 (40 – 160)	0	0.9253	0.9579	0.0326
Mean IAAs (nU/ml)	382 ± 555	382 ± 555	0	346 ± 436	346 ± 436	0	0.4910	0.4910	0
GAD antibodies:									
Positive	144 (77.8)			136 (56.4)					
Negative	41 (22.1)			50 (43.5)					
ICA-512 antibodies									
Positive	97 (52.4)			97 (52.1)					
Negative	88 (47.5)			89 (47.8)					
Micro IAA									
Positive	39 (29.3)			28 (19.4)					
Negative	94 (70.6)			116 (80.5)					

HbA_{1c} (%)	5.35 ± 0.39	5.35 ± 0.39	0	5.33 ± 0.34	5.33 ± 0.34	0	0.5949	0.5949	0
C-peptide area under the curve									
During intravenous glucose tolerance test	34.8 (15.6)			35.1 (16.7)			0.8800		
During oral glucose tolerance test	502.5 (201.1)			502.1 (207.2)			0.9858		
During mixed meal tolerance test	383.1 (172.4)			381.0 (183.8)			0.9102		

Data are means ± SD, n(%), or mean (interquartile range).

* Result is probably a match. The published result is likely a typo.

Table 2: Comparison of Values of Baseline Characteristics of subjects by IAA status Computed in Integrity Check to Reference Article Values

Variable	Not confirmed IAA ≥ 80			Confirmed IAA ≥ 80			P-value		
	Skyler et al	Integrity check	Diff	Skyler et al	Integrity check	Diff	Skyler et al	Integrity check	Diff
N	109	109	0	263	263	0			
Median age	13.0 (9 – 18)	13.7 (9 – 18)	0.7	9.0 (6 – 12)	9.7 (6 – 12)	0.7	0.0000	0.0000	0
Average FPIR (μU/ml)	172.1 ± 73.1	172.1 ± 73.1	0	155.3 ± 91.4	155.3 ± 91.4		0.0878	0.0878	0
Race:							0.9246	0.9741	0.0495
White	96 (88.0)	96 (88.1)	0 (0.1)	231 (87.8)	231 (87.8)	0			
African American	2 (1.8)	2 (1.8)	0	5 (1.9)	5 (1.9)	0			
Hispanic	7 (6.4)	7 (6.4)	0	15 (5.7)	15 (5.7)	0			
Other	4 (3.6)	4 (3.7)	0 (0.1)	12 (4.5)	12 (4.0)	0 (0.1)			
Sex:							0.0445	0.0445	0
Male	57 (52.2)	57 (52.3)	0 (0.1)	167 (63.5)	167 (63.5)	0			
Female	52 (47.7)	52 (47.7)	0	96 (36.5)	96 (36.5)	0			
Relationship to index patient w/ diabetes							0.1649	0.1549	0.01
Sibling	58 (53.2)	58 (53.2)	0	162 (61.6)	162 (61.6)	0			
Offspring	31 (28.4)	31 (28.4)	0	71 (27.0)	72 (27.4)	1 (0.4)			
Parent	9 (8.2)	9 (8.3)	0 (0.1)	9 (3.4)	9 (3.4)	0			
Second degree	11 (10.0)	11 (10.1)	0 (0.1)	21 (7.9)	20 (7.6)	1 (0.3)			
Antibody levels:									
Median ICAs (JDF units)	40 (20 – 160)	40 (20 – 160)	0	80 (40 – 320)	80 (40 – 320)	0	0.0001	0.0001	0
Mean IAAs (nU/ml)	72.0 ± 72.3	72.0 ± 72.3	0	485.2 ± 547.5	485.2 ± 547.5	0	0.0000	0.0000	0
GAD antibodies:							0.0461		
Positive	74 (68.5)			206 (78.3)					
Negative	34 (31.4)			57 (21.6)					
ICA-512 antibodies							0.0043		
Positive	44 (40.7)			150 (57.0)					
Negative	64 (59.2)			113 (42.9)					
Micro IAA							0.0000		
Positive	4 (5.0)			63 (31.9)					
Negative	76 (95.0)			134 (68.0)					
HbA_{1c} (%)	5.33 ± 0.37	5.33 ± 0.37	0	5.35 ± 0.36	5.35 ± 0.36	0	0.6112	0.6112	0

C-peptide area under the curve									
During intravenous glucose tolerance test	40.1 (16.7)			32.8 (15.4)			0.0001		
During oral glucose tolerance test	563.9 (225.0)			476.6 (189.1)			0.0002		
During mixed meal tolerance test	443.2 (183.3)			365.2 (169.5)			0.0000		

Data are means ± SD, n(%), or mean (interquartile range).

Additional explanation of how these variables were defined would be necessary in order to perform a more thorough integrity check of the data. The DCC has agreed to participate in a meeting to discuss these issues, but has also stated that all published results were independently confirmed.

Notes

1. Analysis was done using the copy of the database provided to the NIDDK Data Repository. The SAS datasets used were created on July 27, 2005. 4 of the 31 datasets are examined in this analysis. Of the 27 datasets not examined, 5 of them are related to additional reviews of study data and/or an ancillary study.

2. The SAS datasets provided to the NIDDK Data Repository are in an archival format. In order to use SAS Viewer, limit CPU resources and increase performance when using these datasets, they must be converted back to an un-archived state. One method to do this is via PROC MIGRATE, as follows:

```
/* Location of Archived DPT-1 SAS Data Files */
LIBNAME OLD 'W:\Databases\DPT-1 data\DPT-1_7-27-2005\SAS Data Files';

/* Location for Un-archived DPT-1 SAS Data Files */
LIBNAME NEW 'W:\Project data files\DPT-1';

/* Migrate the datasets */
PROC MIGRATE IN=OLD OUT=NEW; RUN;
```

   Un-archived versions of all the archived datasets in the 'OLD' location will then be created in the 'NEW' location.

Reference

1. The Diabetes Prevention Trial – Type 1 Study Group (Skyler et al). Effects of Oral Insulin in Relatives of Patients With Type 1 Diabetes. **Diabetes Care**, Volume 28, Number 5, May 2005, pages 1068-1076.

ATTACHMENT 1

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

Full Text of Article

The Diabetes Prevention Trial – Type 1 Study Group. Effects Of Oral Insulin in Relatives of Patients with Type 1 Diabetes. Diabetes Care, May 2005;28(5):1068-1076.

ATTACHMENT 2

SAS Code for Baseline Tabulations from DPT-1 Datasets in the NIDDK Repository

```
/*******************************************************************************/
/*
/* Program: R:\Norma\DPT-1\DCare\getdata.sas
/* Author:  Norma Pugh
/* Date:    12 April 06
/* Purpose: Datasets for Diabetes Care article replication.
/*
/*******************************************************************************/
/* Location of NIDDK Repository SAS files */
libname data 'r:\Norma\DPT-1\MigratedData';

/* Location of SAS files created for this replication */
libname out  'r:\Norma\DPT-1\DCare';

/* Randomized subjects */
data out.subjects;
 set data.subject;
 if treatment in('Oral Insulin','Oral Placebo');
 randage=age_at_randomization/365.25;
 keep maskid treatment race gender age_at_randomization randage;
run;

/* Test datasets: First-phase insulin response, Glucose tolerance, ICA, IAA */
data out.fpir(keep=maskid age_at_draw testname ins1 ins3)
     out.ica(keep=maskid age_at_draw testname result outcome)
     out.iaa(keep=maskid age_at_draw testname result outcome);
     out.hba1c(keep=maskid age_at_draw testname result);
     out.peptide(keep=maskid age_at_draw testname pepm10 pepm4 pep0 pep1 pep10 pep120 pep3 pep30
pep5 pep60 pep7 pep90);
 set data.test(keep=maskid age_at_draw testname result outcome ins1 ins3 pepm10 pepm4 pep0 pep1
pep10 pep120 pep3 pep30 pep5 pep60 pep7 pep90);
 if testname in('AB-IVGTT','CO-IVGTT') then output out.fpir;
 if testname='ICA' then output out.ica;
 if testname='IAA' then output out.iaa;
 if testname='HBA1C' then output out.hba1c;
 if testname in('CO-IVGTT','MMTT','OGTT') then output out.peptide;
run;
```

```
/*****************************************************************************/
/*
/* Program: R:\Norma\DPT-1\DCare\table1.sas
/* Author:  Norma Pugh
/* Date:    12 April 06
/* Purpose: Diabetes Care baseline replication (Table 1).
/*
/*****************************************************************************/
/* Location of SAS files created for this replication */
libname data 'R:\Norma\DPT-1\DCare';

/* Location of NIDDK Repository SAS files */
libname master 'R:\Norma\DPT-1\MigratedData';


/***********************************/
/* Age: median, interquartile range */
/***********************************/
proc sort data=data.subjects out=subjects; by treatment; run;

proc univariate data=subjects; by treatment; var randage; title'Age Results'; run;

data subjects; set subjects; newage=log(randage); run; *** no difference ***;

proc npar1way data=subjects;
 class treatment;
 var randage;
run;


/*************************************************************************/
/* First phase insulin response - Add insulin values at 1 and 3 minutes */
/*************************************************************************/
proc sort data=data.fpir out=fpir; by maskid age_at_draw; run;
proc sort data=subjects; by maskid; run;

data fpir;
 merge fpir subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

data fpir;
 set fpir;
 by maskid age_at_draw;
 if last.maskid;
                                            /* Keep final baseline record */
 if substr(ins1,1,1)='<' then I1=0; else I1=ins1+0;  /* Convert to numeric. Recode '<' values to
0. */
 if substr(ins3,1,1)='<' then I3=0; else I3=ins3+0;
 fpir=I1+I3;
run;

proc sort data=fpir; by treatment; run;

proc glm data=fpir;
 class treatment;
 model fpir=treatment / ss3;
 means treatment;
 title'First-phase insulin response Results';
run;
```

```
/*****************************/
/* Race & Gender frequencies */
/*****************************/
data subjects;
 set subjects;
 length finalrace $16.;
 if race='White' then finalrace='White';
 else if race='Black, not Hispanic' then finalrace='African American';
 else if race='Hispanic' then finalrace='Hispanic';
 else if race in('Asian/Pacific Islander','Other','Unknown') then finalrace='Other';
 else finalrace='problem';
run;

data test; set subjects;
 if finalrace='White' then test1='y';
 else if finalrace^='White' then test1='n';
run;

proc sort data=subjects; by treatment; run;
proc freq data=subjects; tables treatment*finalrace / exact; title'Race counts & p-value'; run;
proc freq data=test; tables treatment*test1 / exact; title'test race p-value'; run;
proc freq data=subjects; tables treatment*gender / chisq; title'Gender counts & p-value'; run;


/*********************************/
/* Relationship to index patient */
/*********************************/
proc sort data=master.relative out=relative(keep=maskid reltype qual); by maskid; run;
proc sort data=subjects; by maskid; run;

data relation;
 merge relative subjects(in=keep);
 by maskid;
 if keep;
run;

proc freq data=relation; tables reltype; title'Check values of reltype';  run;

/* NOTE: 2 subjects are classified as reltype='X' (other).  Currently counted as 2nd degree. */
data relation;
 set relation;
 if reltype in('B','IT','SIB','SIS','TWIN') then relation='1_sibling';     /* Categorize
relatives */
 else if reltype in('F','M') then relation='2_parent';
 else if reltype in('D','SON') then relation='3_offsprg';
 else if reltype
in('A','C','FC','GA','GF','GGF','GGM','GM','GP','GU','MA','MC','MGA','MGF','MGM','MU','NE','PA','
PC','PGF','PGM','PU','SC','U','X') then relation='4_2nd_deg';
run;

proc sort data=relation; by maskid relation; run;                                        /* Keep
closest proband */
data relation; set relation; by maskid relation; if first.maskid; run;

proc freq data=relation; tables treatment*relation / chisq; title'Relationship to index patient
Results'; run;

/* TEST: Would p-value match if table values matched exactly? - ANSWER: YES! */
/*
data test;
 input trt $ rel $ count;
 cards;
```

```
 1 1 112
 1 2  49
 1 3  11
 1 4  14
 2 1 108
 2 2  53
 2 3   7
 2 4  18
 ;
run;
title'Check p-value if table values matched exactly';
proc freq; weight count; tables trt*rel / chisq;
*/

/****************/
/* Baseline ICA */
/****************/
proc sort data=data.ica out=ica; by maskid age_at_draw; run;

data ica;
 merge ica subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

data ica;
 set ica;
 by maskid age_at_draw;
 if last.maskid;
                                            /* Keep final baseline record */
 icavalue=result+0;
                                            /* Convert to numeric */
run;

proc sort data=ica; by treatment; run;

proc univariate data=ica; by treatment; var icavalue; title'ICA Results'; run;

proc npar1way data=ica;
 class treatment;
 var icavalue;
run;

/* ICA-512 Antibodies */
proc freq data=ica; tables treatment*outcome / chisq; title'ICA-512 Results'; run;


/****************/
/* Baseline IAA */
/****************/
proc sort data=data.iaa out=iaa; by maskid age_at_draw; run;

data iaa;
 merge iaa subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

data iaa;
 set iaa;
 by maskid age_at_draw;
 if last.maskid;
                                            /* Keep final baseline record */
```

```
 iaavalue=result+0;
                                          /* Convert to numeric */
run;

proc sort data=iaa; by treatment; run;

proc glm data=iaa;
 class treatment;
 model iaavalue=treatment / ss3;
 means treatment;
 title'IAA Results';
run;

/* Micro IAA */
proc freq data=iaa; tables treatment*outcome / chisq; title'Micro IAA Results'; run;


/*********/
/* HbA1c */
/*********/
proc sort data=data.hba1c out=hba1c; by maskid age_at_draw; run;

data hba1c;
 merge hba1c subjects(in=keep);
 by maskid;
 if keep and first.maskid;    /* Keep first HbA1c measurement - only 1 "true" baseline record
(i.e., on/before randomization date) */
 hbavalue=result+0;                                /* Convert to numeric */
run;

proc sort data=hba1c; by treatment; run;

proc glm data=hba1c;
 class treatment;
 model hbavalue=treatment / ss3;
 means treatment;
 title'HbA1c Results';
run;
```

```
/*******************************************************************************/
/*
/* Program: R:\Norma\DPT-1\DCare\table2.sas
/* Author:  Norma Pugh
/* Date:    12 April 06
/* Purpose: Diabetes Care baseline replication (Table 2).
/*
/*******************************************************************************/
/* Location of SAS files created for this replication */
libname data 'R:\Norma\DPT-1\DCare';

/* Location of NIDDK Repository SAS files */
libname master 'R:\Norma\DPT-1\MigratedData';


/*****************/
/* Stratification */
/*****************/
proc sort data=data.subjects out=subjects; by maskid; run;

proc sort data=data.iaa out=iaa; by maskid age_at_draw; run;

data iaa;
 merge iaa subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

data subjects(drop=testname outcome age_at_draw treatment temp result);
 set iaa;
 by maskid;
 retain temp 0;
 if first.maskid then do;
  if result>=80 then temp=1;
  if result<80 then temp=0;
 end;
 else if result>=80 then temp+1;
 if last.maskid;
 if temp>=2 then group='y_confirm';
  else if 0<=temp<=1 then group='n_confirm';
run;


/**********************************/
/* Age: median, interquartile range */
/**********************************/
proc sort data=subjects; by group; run;

proc univariate data=subjects; by group; var randage; title'Age Results'; run;

data subjects; set subjects; newage=log(randage); run; *** no difference ***;

proc npar1way data=subjects;
 class group;
 var randage;
run;


/*********************************************************************/
/* First phase insulin response - Add insulin values at 1 and 3 minutes */
/*********************************************************************/
proc sort data=data.fpir out=fpir; by maskid age_at_draw; run;
proc sort data=subjects; by maskid; run;
```

```
data fpir;
 merge fpir subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

proc sort data=fpir; by maskid age_at_draw; run;

data fpir;
 set fpir;
 by maskid age_at_draw;
 if last.maskid;
                                                /* Keep final baseline record */
 if substr(ins1,1,1)='<' then I1=0; else I1=ins1+0;  /* Convert to numeric. Recode '<' values to
0. */
 if substr(ins3,1,1)='<' then I3=0; else I3=ins3+0;
 fpir=I1+I3;
run;

proc sort data=fpir; by group; run;

proc glm data=fpir;
 class group;
 model fpir=group / ss3;
 means group;
 title'First-phase insulin response Results';
run;


/*****************************/
/* Race & Gender frequencies */
/*****************************/
data subjects;
 set subjects;
 length finalrace $16.;
 if race='White' then finalrace='White';
 else if race='Black, not Hispanic' then finalrace='African American';
 else if race='Hispanic' then finalrace='Hispanic';
 else if race in('Asian/Pacific Islander','Other','Unknown') then finalrace='Other';
 else finalrace='problem';
run;

data test; set subjects;
 if finalrace='White' then test1='y';
 else if finalrace^='White' then test1='n';
run;

proc sort data=subjects; by group; run;
proc freq data=subjects; tables group*finalrace / exact; title'Race counts & p-value'; run;
proc freq data=test; tables group*test1 / exact; title'test race p-value'; run;
proc freq data=subjects; tables group*gender / chisq; title'Gender counts & p-value'; run;


/*********************************/
/* Relationship to index patient */
/*********************************/
proc sort data=master.relative out=relative(keep=maskid reltype qual); by maskid; run;
proc sort data=subjects; by maskid; run;

data relation;
 merge relative subjects(in=keep);
 by maskid;
```

```
 if keep;
run;

proc freq data=relation; tables reltype; title'Check values of reltype';  run;

/* NOTE: 2 subjects are classified as reltype='X' (other).  Currently counted as 2nd degree. */
data relation;
 set relation;
 if reltype in('B','IT','SIB','SIS','TWIN') then relation='1_sibling';     /* Categorize
relatives */
 else if reltype in('F','M') then relation='2_parent';
 else if reltype in('D','SON') then relation='3_offsprg';
 else if reltype
in('A','C','FC','GA','GF','GGF','GGM','GM','GP','GU','MA','MC','MGA','MGF','MGM','MU','NE','PA','
PC','PGF','PGM','PU','SC','U','X') then relation='4_2nd_deg';
run;

proc sort data=relation; by maskid relation; run;
       /* Keep closest proband */
data relation; set relation; by maskid relation; if first.maskid; run;

proc freq data=relation; tables group*relation / chisq; title'Relationship to index patient
Results'; run;


/****************/
/* Baseline ICA */
/****************/
proc sort data=data.ica out=ica; by maskid age_at_draw; run;

data ica;
 merge ica subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

data ica;
 set ica;
 by maskid age_at_draw;
 if last.maskid;
                                           /* Keep final baseline record */
 icavalue=result+0;
                                           /* Convert to numeric */
run;

proc sort data=ica; by group; run;

proc univariate data=ica; by group; var icavalue; title'ICA Results'; run;

proc npar1way data=ica;
 class group;
 var icavalue;
run;

/* ICA-512 Antibodies */
proc freq data=ica; tables group*outcome / chisq; title'ICA-512 Results'; run;


/****************/
/* Baseline IAA */
/****************/
proc sort data=data.iaa out=iaa; by maskid age_at_draw; run;
```

```
data iaa;
 merge iaa subjects(in=keep);
 by maskid;
 if keep & age_at_draw<=age_at_randomization;       /* Keep all baseline records */
run;

data iaa;
 set iaa;
 by maskid age_at_draw;
 if last.maskid;
                                            /* Keep final baseline record */
 iaavalue=result+0;
                                            /* Convert to numeric */
run;

proc sort data=iaa; by group; run;

proc glm data=iaa;
 class group;
 model iaavalue=group / ss3;
 means group;
 title'IAA Results';
run;

/* Micro IAA */
proc freq data=iaa; tables group*outcome / chisq; title'Micro IAA Results'; run;


/*********/
/* HbA1c */
/*********/
proc sort data=data.hba1c out=hba1c; by maskid age_at_draw; run;

data hba1c;
 merge hba1c subjects(in=keep);
 by maskid;
 if keep and first.maskid;    /* Keep first HbA1c measurement - only 1 "true" baseline record
(i.e., on/before randomization date) */
 hbavalue=result+0;                                 /* Convert to numeric */
run;

proc sort data=hba1c; by group; run;

proc glm data=hba1c;
 class group;
 model hbavalue=group / ss3;
 means group;
 title'HbA1c Results';
run;
```