Dataset Integrity Check for the Boston Area Community Health (BACH) III Data Files

Prepared by Allyson Mateja
IMS, Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD 20705
May 25, 2016

Contents

1 Standard Disclaimer

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 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. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless NIDDK Repository staff 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 footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

2 Study Background

The Boston Area Community Health cohort is a multi-stage 1:1:1 stratified random sample of self-identified African American, Hispanic, and white adults from three Boston inner city areas. The goal of the study was to test among diabetes-free urban community-dwelling adults the hypothesis that the proportion of African genetic ancestry is positive associated with glycaemia, after accounting for other continental ancestry proportions, BMI, and socioeconomic status. It was found that a greater proportion of African genetic ancestry is independently associated with higher fasting glucose levels in a non-diabetic community-based cohort, even accounting for other ancestry proportions, obesity, and socioeconomic status. The results suggest that differences between African-Americans and whites in type 2 diabetes risk may include genetically mediated differences in glucose homeostasis.

3 Archived Datasets

The SAS data files, as provided by the Data Coordinating Center (DCC), are located in the data package. For this replication, variables were taken from the "bach3publicuse.sas7bdat" data file.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Meigs, et al. in Diabetologia in 2014 [1]. To verify the integrity of the datasets, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Study cohort characteristics overall and by self-reported race/ethnicity, Table A lists the variables that can be used in the replication. Table B compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are almost an exact match.

Note that a multiple imputation (MI) algorithm was used to obtain the results in the manuscript. This dataset integrity check was performed by comparing the published results to values computed using the 15 MI datasets provided by the study. However, these 15 MI datasets are not included in the data package.

6 Conclusions

The NIDDK repository is confident that the BACH III data files to be distributed are a true copy of the manuscript data.

7 References

[1] Meigs, J.B., Grant, R.W., Piccolo, R., Lopez, L., Florez, J.C., Porneala, B., Marceau, L., McKinlay, J.B. Association of African genetic ancestry with fasting glucose and HbA1c levels in non-diabetic individuals: the Boston Area Community Helath (BACH) Prediabetes Study. Diabetologia (2014) 57: 1850-1858.

Table A: Variables used to replicate Table 1: Study cohort characteristics overall and by self-reported race/ethnicity

Table Variable	**Dataset Variable**
Race/ethnicity	re
Age, years	agegrp_2
Sex	gender
Income, US$	incomecat_2
Education	degree_2
Occupation	occgrpalt_2
Private insurance	inspr_2
Public insurance	inspu_2
BMI, kg/m²	bmicat_2
Systolic blood pressure, mmHg	sbp_2
Diastolic blood pressure, mmHg	dbp_2
LDL-cholesterol, mmol/l	ldl_cholesterol_2 * 0.02586
Fasting glucose, mmol/l	hemocue_2/18
HbA1c, %	hemoglobin_a1c_2
HbA1c, mmol/mol	10.929*(hemoglobin_a1c_2-2.15)
African ancestry, %	aimsperafrcn_2*100
Native American ancestry, %	aimspernamer_2*100
European ancestry, %	aimspereuro_2*100

Table B: Comparison of values computed in integrity check to reference article Table 1 values

Characteristic	All Manuscript (N=1,387)	All DSIC (N=1,387)	Diff. (N=0)	African-American Manuscript (N=379)	African-American DSIC (N=379)	Diff. (N=0)
Age, years						
34-44	281 (30.38)	281 (30.38)	0 (0)	71 (32.31)	71 (32.31)	0 (0)
34-54	421 (28.31)	421 (28.31)	0 (0)	136 (34.86)	136 (34.86)	0 (0)
55-64	372 (18.25)	372 (18.25)	0 (0)	86 (15.78)	86 (15.78)	0 (0)
65-74	209 (12.28)	209 (12.28)	0 (0)	65 (12.06)	65 (12.06)	0 (0)
75-87	104 (10.79)	104 (10.79)	0 (0)	21 (4.99)	21 (4.99)	0 (0)
Sex (% male)	519 (44.95)	515 (44.95)	0 (0)	135 (39.98)	135 (39.98)	0 (0)
Income, US$						
<20,000	455 (22.53)	455 (22.53)	0 (0)	141 (33.44)	141 (33.44)	0 (0)
20,000-49,999	405 (25.65)	405 (25.65)	0 (0)	129 (34.80)	129 (34.80)	0 (0)
≥50,000	527 (51.83)	527 (51.83)	0 (0)	109 (31.77)	109 (31.77)	0 (0)
Education						
Less than high school	174 (6.48)	174 (6.48)	0 (0)	41 (8.43)	41 (8.43)	0 (0)
High school or equivalent	397 (23.12)	397 (23.12)	0 (0)	144 (38.53)	144 (38.53)	0 (0)
Some college	290 (20.65)	290 (20.65)	0 (0)	106 (32.59)	106 (32.59)	0 (0)
College or more	526 (49.75)	526 (49.75)	0 (0)	88 (20.45)	88 (20.45)	0 (0)
Occupation						
Professional, managerial, sales, and office	794 (67.97)	794 (67.97)	0 (0)	213 (59.32)	213 (59.32)	0 (0)
Service	333 (17.03)	333 (17.03)	0 (0)	93 (22.86)	93 (22.86)	0 (0)
Manual labor	187 (11.76)	187 (11.76)	0 (0)	63 (16.27)	63 (16.27)	0 (0)
Never worked	74 (3.24)	74 (3.24)	0 (0)	10 (1.55)	10 (1.55)	0 (0)
Private insurance	807 (70.27)	807 (70.27)	0 (0)	199 (55.57)	199 (55.57)	0 (0)
Public insurance	676 (40.67)	676 (48.67)	0 (0)	206 (51.37)	206 (51.37)	0 (0)
BMI, kg/m²						
Normal (<25)	337 (26.30)	337 (26.30)	0 (0)	72 (16.24)	72 (16.24)	0 (0)
Overweight (25-29)	515 (39.23)	515 (39.23)	0 (0)	132 (34.47)	132 (34.47)	0 (0)
Obese (≥30)	535 (34.46)	535 (34.46)	0 (0)	175 (49.29)	175 (49.29)	0 (0)
Systolic blood pressure, mmHg	128 (0.74)	128 (0.74)	0 (0)	133 (1.49)	133 (1.49)	0 (0)
Diastolic blood pressure, mmHg	80 (0.43)	80 (0.43)	0 (0)	84 (0.84)	84 (0.84)	0 (0)
LDL-cholesterol, mmol/l	2.89 (0.31)	2.89 (0.03)	0 (0.28)	3.01 (0.06)	3.01 (0.06)	0 (0)
Fasting glucose, mmol/l	5.63 (0.03)	5.63 (0.03)	0 (0)	5.72 (0.04)	5.72 (0.04)	0 (0)

Characteristic	All Manuscript (N=1,387)	All DSIC (N=1,387)	Diff. (N=0)	African-American Manuscript (N=379)	African-American DSIC (N=379)	Diff. (N=0)
HbA1c, %	5.53 (0.01)	5.53 (0.01)	0 (0)	5.68 (0.03)	5.68 (0.03)	0 (0)
HbA1c, mmol/mol	36.9 (0.14)	37.0 (0.15)	0.1 (0.01)	38.6 (0.32)	38.6 (0.35)	0 (0.03)
African ancestry, %	28.3 (1.41)	28.3 (1.41)	0 (0)	78.2 (1.12)	78.2 (1.12)	0 (0)
Native American ancestry, %	7.58 (0.45)	7.58 (0.45)	0 (0)	5.95 (0.48)	5.95 (0.48)	0 (0)
European ancestry, %	64.2 (1.50)	64.2 (1.50)	0 (0)	15.9 (1.06)	15.9 (1.06)	0 (0)

Characteristic	Hispanic Manuscript (N=411)	Hispanic DSIC (N=411)	Diff. (N=0)	White Manuscript (N=597)	White DSIC (N=597)	Diff. (N=0)
Age, years						
34-44	99 (45.61)	99 (45.61)	0 (0)	111 (26.60)	111 (26.60)	0 (0)
34-54	147 (30.47)	147 (30.47)	0 (0)	138 (25.35)	138 (25.35)	0 (0)
55-64	118 (16.73)	118 (16.73)	0 (0)	168 (19.50)	168 (19.50)	0 (0)
65-74	34 (4.62)	34 (4.62)	0 (0)	110 (13.89)	110 (13.89)	0 (0)
75-87	13 (2.57)	13 (2.57)	0 (0)	70 (14.67)	70 (14.67)	0 (0)
Sex (% male)	143 (45.86)	143 (45.86)	0 (0)	241 (46.69)	241 (46.69)	0 (0)
Income, US$						
<20,000	205 (39.51)	205 (39.51)	0 (0)	109 (14.92)	109 (14.92)	0 (0)
20,000-49,999	146 (36.61)	146 (36.61)	0 (0)	129 (19.92)	129 (19.92)	0 (0)
≥50,000	59 (23.89)	59 (23.89)	0 (0)	358 (65.16)	358 (65.16)	0 (0)
Education						
Less than high school	117 (22.24)	117 (22.24)	0 (0)	16 (2.57)	16 (2.57)	0 (0)
High school or equivalent	156 (37.46)	156 (37.46)	0 (0)	96 (14.29)	96 (14.29)	0 (0)
Some college	76 (18.11)	76 (18.11)	0 (0)	109 (16.54)	109 (16.54)	0 (0)
College or more	62 (22.18)	62 (22.18)	0 (0)	376 (66.59)	376 (66.59)	0 (0)
Occupation						
Professional, managerial, sales, and office	119 (40.42)	119 (40.42)	0 (0)	463 (76.81)	464 (76.81)	0 (0)
Service	171 (32.06)	171 (32.06)	0 (0)	68 (11.78)	68 (11.78)	0 (0)
Manual labor	71 (17.61)	71 (17.61)	0 (0)	52 (8.84)	52 (8.84)	0 (0)
Never worked	50 (9.91)	50 (9.91)	0 (0)	14 (2.57)	14 (2.57)	0 (0)
Private insurance	157 (47.80)	157 (47.80)	0 (0)	451 (80.44)	451 (80.44)	0 (0)
Public insurance	238 (49.29)	238 (49.29)	0 (0)	232 (34.81)	232 (34.81)	0 (0)
BMI, kg/m²						
Normal (<25)	71 (18.00)	71 (18.00)	0 (0)	194 (31.85)	194 (31.85)	0 (0)

Characteristic	Hispanic Manuscript (N=411)	Hispanic DSIC (N=411)	Diff. (N=0)	White Manuscript (N=597)	White DSIC (N=597)	Diff. (N=0)
Overweight (25-29)	162 (44.20)	162 (44.20)	0 (0)	221 (40.09)	221 (40.09)	0 (0)
Obese (≥30)	178 (37.80)	178 (37.80)	0 (0)	182 (28.06)	182 (28.06)	0 (0)
Systolic blood pressure, mmHg	127 (0.99)	127 (0.99)	0 (0)	127 (0.97)	127 (0.97)	0 (0)
Diastolic blood pressure, mmHg	80 (0.75)	80 (0.75)	0 (0)	78 (0.56)	78 (0.56)	0 (0)
LDL-cholesterol, mmol/l	2.99 (0.06)	2.99 (0.06)	0 (0)	2.82 (0.04)	2.82 (0.04)	0 (0)
Fasting glucose, mmol/l	5.67 (0.05)	5.66 (0.05)	0.01 (0)	5.59 (0.04)	5.59 (0.04)	0 (0)
HbA1c, %	5.57 (0.02)	5.57 (0.02)	0 (0)	5.47 (0.02)	5.47 (0.02)	0 (0)
HbA1c, mmol/mol	37.4 (0.23)	38.4 (0.23)	0 (0)	36.3 (0.17)	36.3 (0.17)	0 (0)
African ancestry, %	30.0 (1.97)	30.0 (1.97)	0 (0)	8.63 (0.54)	8.63 (0.54)	0 (0)
Native American ancestry, %	25.3 (2.15)	25.3 (2.15)	0 (0)	4.68 (0.33)	4.68 (0.33)	0 (0)
European ancestry, %	44.8 (2.05)	44.8 (2.05)	0 (0)	86.7 (0.65)	86.7 (0.65)	0 (0)

All categorical variables are n (%); all continuous variables are mean (SE).

Attachment A: SAS Code

```
**** BACH III DSIC;
**** Programmer: Allyson Mateja;
**** Date: March 22, 2016;

title 'BACH 3 DSIC';
title2 ' ';

libname bachdata '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK';
libname bach3mi1 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi1/';
libname bach3mi2 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi2/';
libname bach3mi3 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi3/';
libname bach3mi4 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi4/';
libname bach3mi5 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi5/';
libname bach3mi6 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi6/';
libname bach3mi7 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi7/';
libname bach3mi8 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi8/';
libname bach3mi9 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi9/';
libname bachmi10 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi10/';
libname bachmi11 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi11/';
libname bachmi12 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi12/';
libname bachmi13 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi13/';
libname bachmi14 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi14/';
libname bachmi15 '/prj/niddk/ims_analysis/BACH3/private_orig_data/BACH 3 NIDDK/BACH 3 NIDDK/m01a_mi15/';

options nofmterr;

proc format;
       value agef 1 = '34-44'
            2 = '45-54'
            3 = '55-64'
            4 = '65-74'
            5 = '75-87';

       value genderf 1 = 'M'
                     2 = 'F';

       value educf 1 = 'Less than high school'
                   2 = 'High school or equivalent'
                   3 = 'Some college'
                   4 = 'College or more';

       value occuf 1 = 'Professional, managerial, sales, and office'
                   2 = 'Service'
                   3 = 'Manual labor'
                   4 = 'Never worked';

       value bmif 1 = 'Normal (< 25)'
                  2 = 'Overweight (25-29)'
                  3 = 'Obese (>= 30)';

data bach;
```

```
        set bachdata.bach3publicuse;

proc contents data = bach;

data mi1;
        set bach3mi1.m01a_mi1;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi2;
        set bach3mi2.m01a_mi2;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi3;
        set bach3mi3.m01a_mi3;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi4;
        set bach3mi4.m01a_mi4;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi5;
        set bach3mi5.m01a_mi5;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
```

```
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi6;
        set bach3mi6.m01a_mi6;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi7;
        set bach3mi7.m01a_mi7;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi8;
        set bach3mi8.m01a_mi8;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi9;
        set bach3mi9.m01a_mi9;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;
```

```
data mi10;
        set bachmi10.m01a_mi10;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi11;
        set bachmi11.m01a_mi11;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi12;
        set bachmi12.m01a_mi12;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi13;
        set bachmi13.m01a_mi13;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi14;
        set bachmi14.m01a_mi14;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
```

```
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

data mi15;
        set bachmi15.m01a_mi15;
        if re = 'B' then race = 1;
        if re = 'H' then race = 2;
        if re = 'W' then race = 3;
        hba1c_mmol = 10.929*(hemoglobin_a1c_2 - 2.15);
        fg_mmol = hemocue_2/18;
        ldl_chol_mmol = ldl_cholesterol_2 * 0.02586;
        aims_percentage_euro = aimspereuro_2*100;
        aims_percentage_afrcn = aimsperafrcn_2*100;
        aims_percentage_namer = aimspernamer_2*100;

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
             nest stratum psu;
             weight wt_m01a;
             subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
             tables race*AGEGRP_2;
             subgroup race agegrp_2;  levels 3 5;
      run;
        title3 'Table 1, Age';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
             nest stratum psu;
             weight wt_m01a;
             subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
             tables race*gender;
             subgroup race gender;  levels 3 2;
      run;
        title3 'Table 1, Gender';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
             nest stratum psu;
             weight wt_m01a;
             subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
             tables race*incomecat_2;
             subgroup race incomecat_2;  levels 3 3;
      run;
        title3 'Table 1, Income';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
             nest stratum psu;
             weight wt_m01a;
             subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
             tables race*degree_2;
             subgroup race degree_2;  levels 3 4;
      run;
        title3 'Table 1, Education';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
             nest stratum psu;
             weight wt_m01a;
             subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
```

```
            tables race*occgrpalt_2;
            subgroup race occgrpalt_2;  levels 3 4;
      run;
       title3 'Table 1, Occupation';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            tables race*inspr_2;
            subgroup race inspr_2;  levels 3 2;
      run;
       title3 'Table 1, Private insurance';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            tables race*inspu_2;
            subgroup race inspu_2;  levels 3 2;
      run;
       title3 'Table 1, Public insurance';

proc crosstab data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            tables race*bmicat_2;
            subgroup race bmicat_2;  levels 3 3;
      run;
       title3 'Table 1, BMI';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var sbp_2;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, Systolic blood pressure';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var dbp_2;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, Diastolic blood pressure';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
```

```
            var ldl_chol_mmol;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, LDL Cholesterol';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var fg_mmol;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, Fasting glucose';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var hemoglobin_a1c_2;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, HbA1c %';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var hba1c_mmol;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, HbA1c, mmol/mol';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var aims_percentage_afrcn;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, Percentage African ancestry';

proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var aims_percentage_namer;
            tables race;
            subgroup race; levels 3;
      run;
       title3 'Table 1, Percentage Native American ancestry';
```

```
proc descript data = mi1 filetype = sas mi_count = 15 design = wr;
            nest stratum psu;
            weight wt_m01a;
            subpopn diabexclude_2 = 2 and hemocueallfast_2 <= 125;
            var aims_percentage_euro;
            tables race;
            subgroup race; levels 3;
      run;
        title3 'Table 1, Percentage European ancestry';
```