Susan M Rogers
May 23, 2008

Data Integrity Check for the Diabetes Prevention Project (DPP) 2008 Full Scale Release

As a partial check of the integrity of the 2008 Full Scale Release of the Diabetes Prevention Program (DPP) data archived in the NIDDK data repository, a series of tabulations was performed to verify that published results from the DPP study can be reproduced using the archived datasets. Several analyses were performed to duplicate results published by the DPP Research Group in the following publication:

> Role of Insulin Secretion and Sensitivity in the Evolution of Type 2 Diabetes in the Diabetes Prevention Program: Effects of lifestyle intervention and metformin (2005) *Diabetes* 54: 2404-14

The full text of the selected articles can be found in Attachment 1. STATA (v10) code for our data integrity checks is included in Attachment 2.

Our replicated analyses produced similar results to the published tabulations. Trends in mean plasma glucose and other metabolic variables were similar to those presented in the published articles. Some discrepancies between the published and replicated analyses may have arisen since only study clinic sites whose IRBs approved the distribution of their data to the NIDDK repository are included in the archived data, hence in some cases it is impossible to match the sample Ns.

The 2008 Full-Scale Data Release: Baseline and Follow-up Data
The DPP Research Group reports results for 3,665 participants in the 2008 Full Scale Data Release who were randomly assigned to one of four treatment arms aimed at preventing type 2 diabetes in high-risk adults: metformin (N=1,027), troglitazone (N=584), intensive lifestyle (N=1,024), or placebo (N=1030). Eligibility criteria for the study included: age 25 years and older, a body mass index of 24 kg/m^2 or higher (22+ if Asian American), a fasting plasma glucose (FPG) level of 95-125 mg/dl (<=125 for American Indians), and a 2-hour plasma glucose concentration of 140-199 mg/dl (see 2008 Data Release Documentation for full details). Eligible participants were identified through a multi-step screening and recruitment process. Enrollment began in 1996 and participants were followed through 2001. Follow-up assessments (quarterly, semi-annual, annual, etc) included various physical measurements, medical history updates, questionnaire administration, medication adherence, and medical testing according to a standard protocol.

In de-identifying the data, all personal identifiers including all dates were removed from the DPP archive. Variables that could possibly identify a particular individual were grouped, e.g., race/ethnicity was recoded into 4 groups (Caucasian, African American, Hispanic and Other), age at baseline was recoded in 5-year groupings with truncation of those <40 and 65+, fasting glucose levels less than 100 at baseline appear as 99, and baseline BMI is provided in two alternative groupings (2 kg/m^2 with truncation of those $\leq$ 26 kg/m^2 and those $\geq$ 42 kg/m^2 and

tertiles of <30, 30 to <35, and ≥ 35 kg/m^2. Only research data are included in the released dataset (screening and post-randomization visits, clinic visits, lifestyle visits, laboratory data). Non-research data and adverse event data are not included.

Data Forms
The NIDDK data repository includes 37 data files – 31 files of data collected on standard forms ('form' files) and 6 files of data not collected on forms ('nonform' files) that include laboratory data, nutrition, quality of well-being, CT-scan, a summary of event variables for diabetes, and a baseline file (treatment assignment, baseline age and BMI group, sex, and race/ethnicity). Several variables are common to all datasets and can be used to link multiple files or to match specific visits across multiple forms. These include RELEASE_ID (unique subject ID), VISIT (baseline, screening, annual visits, interim unscheduled visit, etc), and DAYSRAND (number of days between a visit and randomization). For this verification exercise, data from 3 of the 37 files were accessed.

All form files are named with the prefix DPP_REL and either a S (screening form), F (follow-up form), TR (form for participants randomized to troglitazone), Q (questionnaires), L (lifestyle form), E (events form, e.g. pregnancy), or R (report, e.g. CHD risk status) suffix. One dataset exists for each DPP form and corresponding SAS dataset. Variables and measurement times for all non-form files, e.g., LAB, Nutrient Data, CT Scan Data, Quality of Well Being Data, Baseline Data, and Events Data, are provided in the Data Release Documentation (pages 16-26).

The Diabetes Prevention Program Research Group (2005) Role of Insulin Secretion and Sensitivity in the Evolution of Type 2 Diabetes in the Diabetes Prevention Program: Effects of lifestyle intervention and metformin *Diabetes* 54:2404-2414.

This manuscript evaluates indexes of insulin secretion and sensitivity, the use of metformin or an intensive lifestyle intervention, and their relationship to the development of type 2 diabetes. Only participants randomized to either the metformin, lifestyle, or placebo are included (N=3,234) in the published article; the NIDDK repository includes data on 3,081 assigned to these interventions (N=1,024, lifestyle; N=1,027 metformin; N=1,030 placebo). A copy of the manuscript is included in Attachment 1. STATA programming code to replicate Figures 1 and 2 in the published article are shown in Attachment 2.

Figure 1 plots the mean and SE of the plasma glucose levels during OGTT at baseline and year 1 by treatment group. Only participants who underwent OGTT testing at year 1 are included, i.e., the 61 participants who developed diabetes at 6 months are excluded. Variables from the NIDDK repository used to replicate Figure 1 are shown below.

NIDDK variables used in replication of Figure 1

Figure 1 variable	NIDDK variable used in replication
Treatment assignment	DPP_REL.Basedata, assign
Indicator of diabetes	DPP_REL.EVENTS, diabf
Plasma glucose, fasting	DPP_REL.LAB, g000 [visit==BAS, Y01]
Plasma glucose, 30 min	DPP_REL.LAB, g030 [visit==BAS, Y01]
Plasma glucose, 2 hr	DPP_REL.LAB, g120 [visit==BAS, Y01]

Results from the replicated analyses are shown in **FIGURE 2**. The actual data values (sample Ns and mean plasma glucose levels ± SE at baseline and year 1 by treatment) used to create Figure 3 are shown in **TABLE F**. Data on 61 participants in the repository data who developed diabetes at 6 months and were assigned to one of the three treatments were excluded from the tabulations. A comparison of the published and replicated results indicates similar results with the greatest reduction in glucose at year 1 occurring in the lifestyle group. Decreases in fasting and 30- and 120-min glucose values were also greater in the lifestyle group.

Figure 2 of the published manuscript examines changes in metabolic variables (fasting glucose, fasting insulin, and fasting proinsulin) over time and by treatment assignment. The number of participants decreased over time because of the variable length of time that individuals were in the study. For example, in the published table, data on fasting glucose were available for 3,065 at year 1, 3,015 at 2 years, and 1,910 in year 3. The replicated results are graphed in **FIGURE 3** and data are provided in **TABLE G.** Again, our replicated results resemble closely the published results. In our replicated analyses, data on fasting glucose were available for 2,915 at year 1, 2,865 at year 2.

TABLE F. Calculation of mean plasma glucose during OGTT at baseline and year 1 by treatment assignment as reported in Diabetes 54:2406, 2005 using data from the 2008 DPP Full Scale Data Release in the NIDDK Repository

Tabulations for Published Figure 1. Plasma glucose during OGTT at baseline and year 1 by treatment group.

Mean plasma glucose, Baseline

Treatment assignment	Fasting	Time 30-min	120-min
Placebo	106.9 (0.23)	169.4 (0.76)	164.1 (0.5)
Metformin	107.1 (0.24)	170.3 (0.78)	164.8 (0.5)
Lifestyle	106.9 (0.23)	169.2 (0.78)	164.3 (0.5)
N	3,020	2,981	3,020

Mean plasma glucose, Year 1

Treatment assignment	Fasting	Time 30-min	120-min
Placebo	106.8 (.48)	173.3 (.98)	157.7 (1.38)
Metformin	102.5 (.36)	171.1 (0.9)	156.26 (1.24)
Lifestyle	101.5 (.36)	160.7 (0.86)	140.6 (1.17)
N	2,863	2,831	2,856

Tabulated from 2008 DPP Full Scale Data Release in NIDDK repository, DPP_REL.basedata, DPP_REL.lab, DPP_REL.events
Note: Data are means (± SE). Only participants who underwent OGTT testing at year 1 are included (i.e., excludes 61 participants who developed diabetes within 6 months, 10 in the lifestyle group, 12 in the metformin group, and 39 assigned to the placebo).

TABLE G. Calculation of metabolic variables over time by treatment assignment as reported in *Diabetes* 54:2405, 2005 using data from the 2008 DPP Full Scale Data Release in the NIDDK Repository

Tabulations for Published Figure 2. Metabolic variables over time by treatment group.

Fasting Glucose (mg/dl)

Treatment assignment	Baseline mean (SE)	Year 1 mean (SE)	Year 2 mean (SE)	Year 3 mean (SE)
Placebo	107.04 (0.23)	101.7 (0.37)	105.3 (0.48)	107.4 (0.6)
Metformin	107.3 (0.24)	102.7 (0.36)	106.0 (0.41)	107.1 (0.56)
Lifestyle	107.4 (0.24)	107.59 (0.49)	109.78 (0.46)	111.6 (0.74)
N	3,081	2,915	2,865	1,817

Fasting Insulin

Treatment assignment	Baseline	Year 1	Year 2	Year 3
Placebo	26.39 (0.49)	21.2 (0.49)	22.3 (0.49)	24.0 (0.6)
Metformin	26.9 (0.46)	23.3 (0.44)	25.2 (0.55)	25.7 (0.65)
Lifestyle	26.4 (0.45)	27.2 (0.56)	27.2 (0.52)	27.2 (0.58)
N	3,078	2,845	2,779	1,724

Fasting Proinsulin (pmol/l)

Treatment assignment	Baseline	Year 1	Year 2	Year 3
Placebo	18.01 (0.4)	12.9 (0.39)	14.2 (0.48)	15.87 (0.7)
Metformin	18.35 (0.46)	13.8 (0.38)	15.3 (0.46)	16.8 (0.7)
Lifestyle	18.11 (0.43)	18.3 (0.54)	18.7 (0.61)	19.16 (0.7)
N	3,075	2,845	2,784	1,727

Note: Data are means (std error)
Tabulated from 2008 DPP Full Scale Data Release in NIDDK repository, DPP_REL.basedata, DPP_REL.lab, DPP_REL.events

FIGURE 3. Tabulation of mean metabolic variables over time by treatment group calculated from the 2008 DPP Full Scale Data Release in the NIDDK repository

FIGURE 2. Tabulation of mean plasma glucose during OGTT at baseline and year one by treatment group calculated from the 2008 DPP Full Scale Data Release in the NIDDK repository
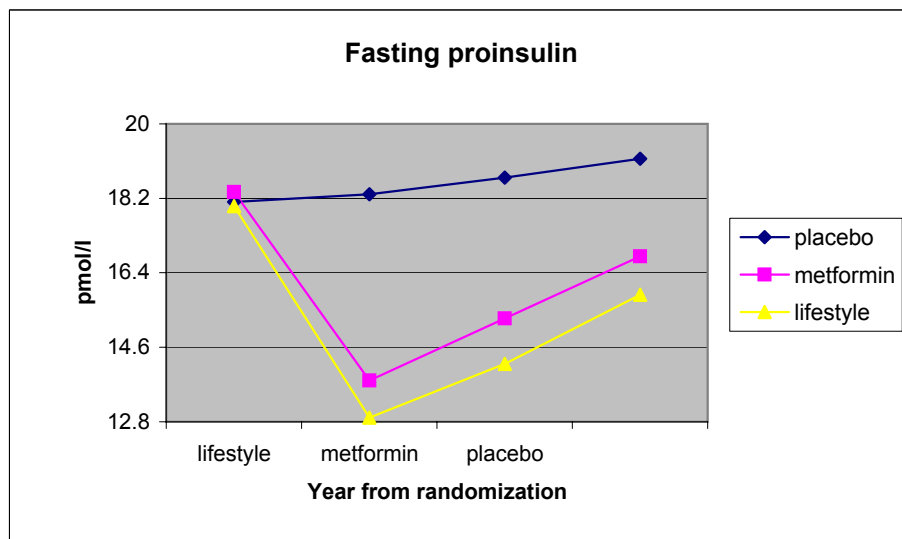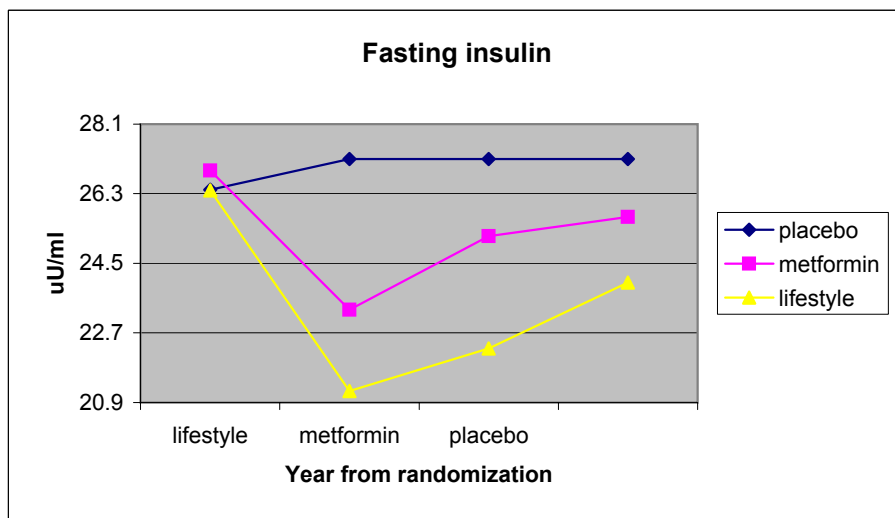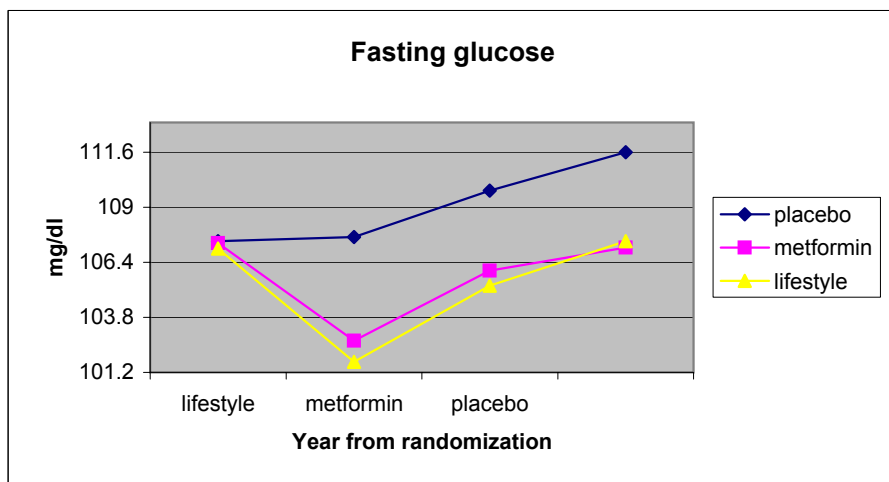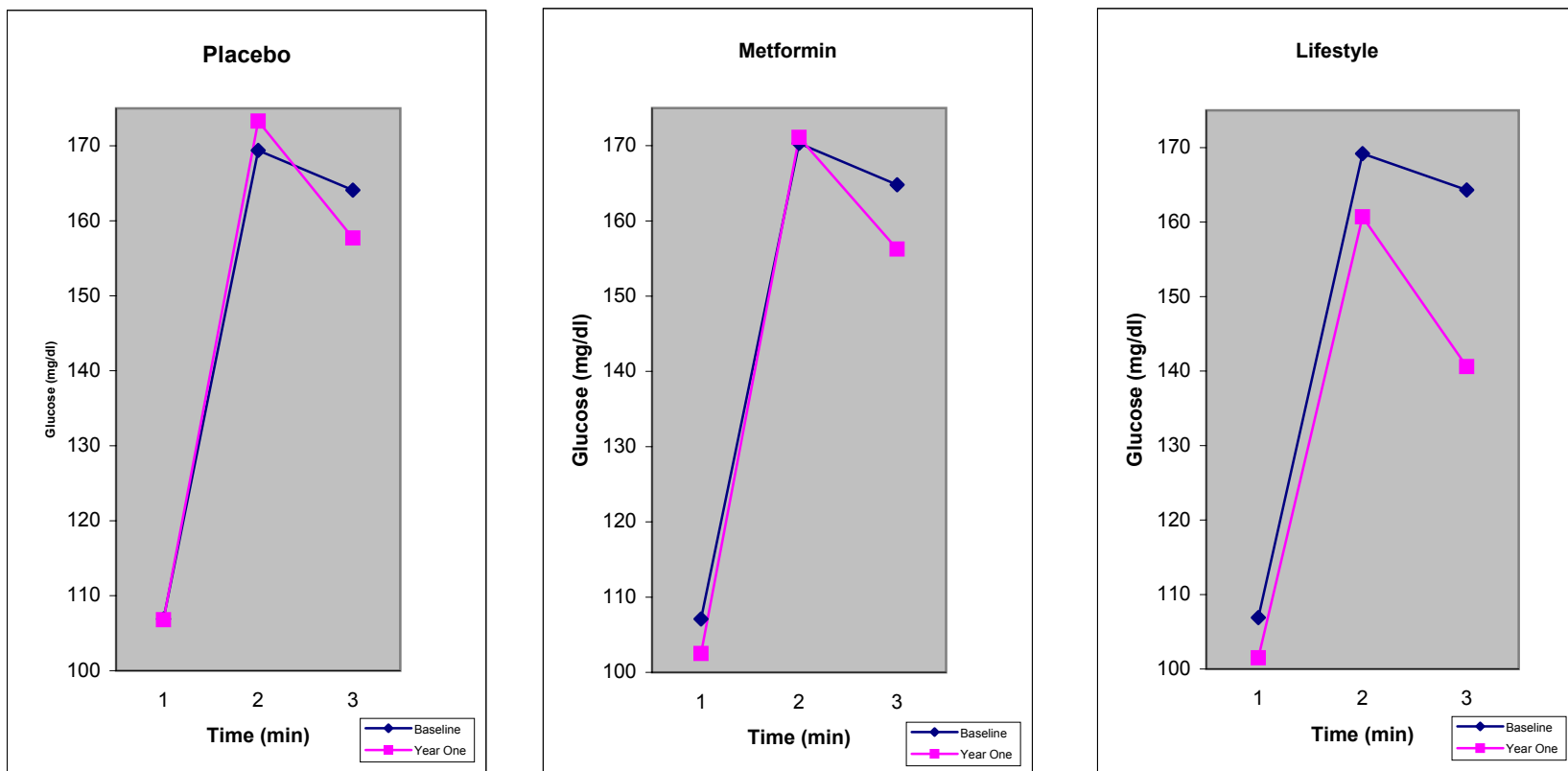
Note: Original figure published in Diabetes Prevention Program Research Group (2005) Role of insulin secretion and sensitivity in the evolution of type 2 diabetes in the Diabetes Prevention Program *Diabetes* 54:2404-2414, Figure 1.

Data shown are means. Data means and SE are shown in TABLE F. Only participants who underwent OGTT testing at year one are included, i.e., those who did not develop diabetes at 6 months.

ATTACHMENT 1

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

The Diabetes Prevention Program Research Group (2005) Role of Insulin Secretion and Sensitivity in the Evolution of Type 2 Diabetes in the Diabetes Prevention Program: Effects of lifestyle intervention and metformin *Diabetes* **54:2404-2414**

ATTACHMENT 2

STATA/SE10 Code for Tabulations of Insulin Secretion and Sensitivity from the DPP Dataset in the NIDDK Repository [*Diabetes* 54: 2404-2414, 2005; Figure 1 & Figure 2]

```
--------------------------------------------------------------------------------
        log:  C:\DPP\analyses\Diab2005_54\insulin.log
   log type:  text
 opened on:  22 May 2008, 09:58:37

. use "C:\DPP\analyses\Diab2005_54\pgdat.dta", clear

. ***tabulations for Diabetes 2005 article, insulin secretion & sensitivity &
developmentof type 2 diabetes
. **Fig 2 fasting glucose, insulin, proinsulin by treatment and time
. **Figure 1 Plasma glucose @ baseline, year 1 by treatment
. ****merge lab-base data and events data
.
. *save c:\DPP\analyses\Diab2005_54\pgdat.dta
.
. /*create randomization variable excludes Rs assigned to Troglitazone
> gen random=.
> replace random=1 if assign=="Lifestyle"
> replace random=2 if assign=="Metformin"
> replace random=3 if assign=="Placebo"
> label define random 1"Lifestyle" 2"Metformin" 3"Placebo"
> label val random random*/
.
. tab random if visit=="BAS"

      random |      Freq.     Percent        Cum.
------------+-----------------------------------
   Lifestyle |      1,024       33.24       33.24
   Metformin |      1,027       33.33       66.57
     Placebo |      1,030       33.43      100.00
------------+-----------------------------------
       Total |      3,081      100.00

.
. *Figure 2. Fasting Glucose
. sort random

. mean g000 if visit=="BAS", over (random)

Mean estimation                       Number of obs   =     3081

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


--------------------------------------------------------------------
        Over |       Mean   Std. Err.      [95% Conf. Interval]
-------------+------------------------------------------------------
g000         |
   Lifestyle |     107.04   .2328172      106.5835     107.4965
   Metformin |   107.2707    .245287      106.7897     107.7516
     Placebo |   107.4194   .2439002      106.9412     107.8976
--------------------------------------------------------------------
.
. *To convert mg/dl of glucose to mmol/l, divide by 18 or multiply by 0.055.
. gen G000R=g000*.055
(4094 missing values generated)
```

```
. mean G000R if visit=="BAS", over (random)

Mean estimation                         Number of obs   =    3081

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


      ------------------------------------------------------------
          Over |      Mean   Std. Err.     [95% Conf. Interval]
      ---------+--------------------------------------------------
      G000R    |
     Lifestyle |  5.887202   .0128049      5.862095    5.912309
     Metformin |  5.899888   .0134908      5.873436     5.92634
       Placebo |  5.908068   .0134145      5.881766     5.93437
      ------------------------------------------------------------


.
. *insulin
. mean i000 if visit=="BAS", over (random)

Mean estimation                         Number of obs   =    3078

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


      ------------------------------------------------------------
          Over |      Mean   Std. Err.     [95% Conf. Interval]
      ---------+--------------------------------------------------
      i000     |
     Lifestyle |  26.39404   .4862537      25.44062    27.34745
     Metformin |  26.91462   .4638347      26.00516    27.82408
       Placebo |  26.41429   .4529247      25.52622    27.30235
      ------------------------------------------------------------


.
. *glucose conversion uU/mL to pmol/L multiply by 6
. *gen i000R=i000*6

. mean i000R if visit=="BAS", over (random)

Mean estimation                         Number of obs   =    3078

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


      ------------------------------------------------------------
          Over |      Mean   Std. Err.     [95% Conf. Interval]
      ---------+--------------------------------------------------
      i000R    |
     Lifestyle |  158.3642   2.917522      152.6437    164.0847
     Metformin |  161.4877   2.783008       156.031    166.9445
       Placebo |  158.4857   2.717548      153.1573    163.8141
      ------------------------------------------------------------
```

```
.**Fasting proinsulin
. mean pin if visit=="BAS", over (random)

Mean estimation                          Number of obs    =     3075

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


-----------------------------------------------------------------
       Over |        Mean    Std. Err.     [95% Conf. Interval]
------------+----------------------------------------------------
pin         |
   Lifestyle |    18.01324    .4148378      17.19985    18.82662
   Metformin |    18.34834    .4593348      17.44771    19.24898
     Placebo |    18.10972     .433001      17.26072    18.95872
-----------------------------------------------------------------


.
. *year one
. mean g000 if visit=="Y01", over (random)

Mean estimation                          Number of obs    =     2915

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


-----------------------------------------------------------------
       Over |        Mean    Std. Err.     [95% Conf. Interval]
------------+----------------------------------------------------
g000        |
   Lifestyle |    101.7446    .3712528      101.0166    102.4725
   Metformin |    102.7265    .3636078      102.0136    103.4395
     Placebo |    107.5949    .4869994        106.64    108.5498
-----------------------------------------------------------------

. mean G000R if visit=="Y01", over (random)

Mean estimation                          Number of obs    =     2915

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


-----------------------------------------------------------------
       Over |        Mean    Std. Err.     [95% Conf. Interval]
------------+----------------------------------------------------
G000R       |
   Lifestyle |    5.595953    .0204189      5.555916     5.63599
   Metformin |    5.649959    .0199984      5.610746    5.689171
     Placebo |    5.917718     .026785      5.865199    5.970237
-----------------------------------------------------------------


.
. mean i000 if visit=="Y01", over (random)
```

```
Mean estimation                           Number of obs    =    2845

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


-----------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+---------------------------------------------------
i000         |
   Lifestyle |   21.21209   .4914253      20.24851    22.17568
   Metformin |   23.32889   .4393037       22.4675    24.19028
     Placebo |   27.17819   .5622545      26.07572    28.28066
-----------------------------------------------------------------

. mean i000R if visit=="Y01", over (random)

Mean estimation                           Number of obs    =    2845

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


-----------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+---------------------------------------------------
i000R        |
   Lifestyle |   127.2726   2.948552       121.491    133.0541
   Metformin |   139.9733   2.635822       134.805    145.1417
     Placebo |   163.0691   3.373527      156.4543    169.6839
-----------------------------------------------------------------


.
. mean pin if visit=="Y01", over (random)

Mean estimation                           Number of obs    =    2845

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


-----------------------------------------------------------------
        Over |       Mean   Std. Err.     [95% Conf. Interval]
-------------+---------------------------------------------------
pin          |
   Lifestyle |   12.93424    .393234      12.16319     13.7053
   Metformin |   13.79587   .3760253      13.05856    14.53318
     Placebo |   18.30992   .5434266      17.24437    19.37547
-----------------------------------------------------------------


.
. *year two
. mean g000 if visit=="Y02", over (random)

Mean estimation                           Number of obs    =    2865
```

```
      Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


-----------------------------------------------------------------
        Over |       Mean    Std. Err.      [95% Conf. Interval]
-------------+---------------------------------------------------
g000         |
   Lifestyle |   105.2918    .4838863       104.343     106.2406
   Metformin |    106.024    .4138411      105.2126     106.8355
     Placebo |   109.7786    .4622115      108.8723     110.6849
-----------------------------------------------------------------

. mean G000R if visit=="Y02", over (random)

Mean estimation                       Number of obs     =     2865

      Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


-----------------------------------------------------------------
        Over |       Mean    Std. Err.      [95% Conf. Interval]
-------------+---------------------------------------------------
G000R        |
   Lifestyle |   5.791047    .0266137      5.738862     5.843231
   Metformin |   5.831322    .0227613      5.786692     5.875952
     Placebo |   6.037822    .0254216      5.987976     6.087669
-----------------------------------------------------------------


.
. mean i000 if visit=="Y02", over (random)

Mean estimation                       Number of obs     =     2779

      Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


-----------------------------------------------------------------
        Over |       Mean    Std. Err.      [95% Conf. Interval]
-------------+---------------------------------------------------
i000         |
   Lifestyle |   22.28402    .4882458      21.32666     23.24138
   Metformin |   25.21963    .5482734      24.14457      26.2947
     Placebo |   27.19076    .5222279      26.16677     28.21476
-----------------------------------------------------------------

. mean i000R if visit=="Y02", over (random)

Mean estimation                       Number of obs     =     2779

      Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


-----------------------------------------------------------------
```

```
        Over |       Mean   Std. Err.      [95% Conf. Interval]
-------------+------------------------------------------------
i000R        |
   Lifestyle |   133.7041   2.929475      127.9599    139.4483
   Metformin |   151.3178   3.289641      144.8674    157.7682
     Placebo |   163.1446   3.133367      157.0006    169.2885
-------------------------------------------------------------


.
. mean pin if visit=="Y02", over (random)

Mean estimation                      Number of obs   =    2784

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


-------------------------------------------------------------
        Over |       Mean   Std. Err.      [95% Conf. Interval]
-------------+------------------------------------------------
pin          |
   Lifestyle |    14.2137   .4754693      13.28139    15.14601
   Metformin |   15.32121    .463341      14.41269    16.22974
     Placebo |   18.73419   .6132639      17.53169    19.93669
-------------------------------------------------------------


.
. *year three
. mean g000 if visit=="Y03", over (random)

Mean estimation                      Number of obs   =    1817

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


-------------------------------------------------------------
        Over |       Mean   Std. Err.      [95% Conf. Interval]
-------------+------------------------------------------------
g000         |
   Lifestyle |   107.3617   .6004648       106.184    108.5393
   Metformin |   107.1223   .5598229      106.0243    108.2202
     Placebo |   111.6371   .7381516      110.1894    113.0848
-------------------------------------------------------------

. mean G000R if visit=="Y03", over (random)

Mean estimation                      Number of obs   =    1817

    Lifestyle: random = Lifestyle
    Metformin: random = Metformin
      Placebo: random = Placebo


-------------------------------------------------------------
        Over |       Mean   Std. Err.      [95% Conf. Interval]
-------------+------------------------------------------------
G000R        |
```

```
    Lifestyle |   5.904892    .0330256         5.84012    5.969664
    Metformin |   5.891725    .0307903        5.831337    5.952113
     Placebo |    6.14004    .0405983        6.060416    6.219665
------------------------------------------------------------


.
. mean i000 if visit=="Y03", over (random)

Mean estimation                          Number of obs     =      1724

     Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


------------------------------------------------------------
        Over |       Mean    Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------
i000         |
    Lifestyle |   24.03175     .647221        22.76233    25.30118
    Metformin |   25.71291    .6491438        24.43972    26.98611
     Placebo |   27.20207    .5801595        26.06417    28.33996
------------------------------------------------------------

. mean i000R if visit=="Y03", over (random)

Mean estimation                          Number of obs     =      1724

     Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


------------------------------------------------------------
        Over |       Mean    Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------
i000R        |
    Lifestyle |   144.1905    3.883326         136.574    151.8071
    Metformin |   154.2775    3.894863        146.6383    161.9166
     Placebo |   163.2124    3.480957         156.385    170.0397
------------------------------------------------------------


.
. mean pin if visit=="Y03", over (random)

Mean estimation                          Number of obs     =      1727

     Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


------------------------------------------------------------
        Over |       Mean    Std. Err.      [95% Conf. Interval]
-------------+----------------------------------------------
pin          |
    Lifestyle |   15.87408    .7116984         14.4782    17.26996
    Metformin |   16.81307    .6831262        15.47322    18.15291
     Placebo |   19.15997    .6977288        17.79148    20.52845
------------------------------------------------------------
```

```
. **Figure 1 Plasma glucose during OGTT
. tab diabf diabv if visit=="BAS"

  indicator |
         of |                   interval for diabetes
   diabetes |         1         2         3         4         5 |     Total
-----------+-------------------------------------------------------+----------
          0 |       109        48        44       282       632 |     2,863
          1 |        69       205        55       206        63 |       802
-----------+-------------------------------------------------------+----------
      Total |       178       253        99       488       695 |     3,665


  indicator |
         of |                   interval for diabetes
   diabetes |         6         7         8         9        10 |     Total
-----------+-------------------------------------------------------+----------
          0 |       543       599       366       239         1 |     2,863
          1 |       115        40        42         7         0 |       802
-----------+-------------------------------------------------------+----------
      Total |       658       639       408       246         1 |     3,665


.
. gen diab6=.
(31815 missing values generated)

. replace diab6=1 if diabf==1 & diabv==1
(746 real changes made)
. replace diab6=0 if diabf==0
(23210 real changes made)
. replace diab6=0 if diabf==1 & diabv >1
(7859 real changes made)
. label var diab6 "diabetes at 6 mos"

. tab diab6 if visit=="BAS"

diabetes at |
      6 mos |      Freq.      Percent        Cum.
------------+-----------------------------------
          0 |      3,596        98.12       98.12
          1 |         69         1.88      100.00
------------+-----------------------------------
      Total |      3,665       100.00


.
. tab diab6 random if visit=="BAS"

  diabetes |               random
  at 6 mos | Lifestyle  Metformin    Placebo |     Total
-----------+---------------------------------+----------
          0 |     1,014      1,015        991 |     3,020
          1 |        10         12         39 |        61
-----------+---------------------------------+----------
      Total |     1,024      1,027      1,030 |     3,081
```

```
.  *exclude Rs who developed diabetes by 6 months
.  keep if diab6 !=1
(746 observations deleted)
.
.  sort random

.  mean g000 if visit=="BAS", over (random)

Mean estimation                     Number of obs    =     3020

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


      ----------------------------------------------------------------
         Over |       Mean    Std. Err.     [95% Conf. Interval]
      --------------+-------------------------------------------------
      g000          |
      Lifestyle |   106.9221   .2311125      106.4689    107.3752
      Metformin |   107.1458   .2440306      106.6673    107.6243
        Placebo |   106.9041   .2325216      106.4482    107.3601
      ----------------------------------------------------------------


.
.  mean g030 if visit=="BAS", over (random)

Mean estimation                     Number of obs    =     2981

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


      ----------------------------------------------------------------
         Over |       Mean    Std. Err.     [95% Conf. Interval]
      --------------+-------------------------------------------------
      g030          |
      Lifestyle |    169.245   .7820863      167.7115    170.7785
      Metformin |    170.255   .7757743      168.7339    171.7761
        Placebo |   169.3848   .7628727       167.889    170.8806
      ----------------------------------------------------------------


.
.  mean g120 if visit=="BAS", over (random)

Mean estimation                     Number of obs    =     3020

      Lifestyle: random = Lifestyle
      Metformin: random = Metformin
        Placebo: random = Placebo


      ----------------------------------------------------------------
         Over |       Mean    Std. Err.     [95% Conf. Interval]
      --------------+-------------------------------------------------
      g120          |
      Lifestyle |   164.2594   .5254502      163.2291    165.2896
      Metformin |   164.7931   .5372644      163.7397    165.8465
        Placebo |   164.1372   .5425549      163.0734     165.201
```

```
---------------------------------------------------------------
.
.
. *year one measurements
. mean g000 if visit=="Y01", over (random)

Mean estimation                     Number of obs    =     2863

     Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


     ---------------------------------------------------------
         Over |       Mean   Std. Err.    [95% Conf. Interval]
     ---------+-----------------------------------------------
     g000     |
     Lifestyle |   101.5311   .3620472     100.8212     102.241
     Metformin |   102.5198   .3585892     101.8167    103.2229
       Placebo |   106.8424   .4767015     105.9077    107.7771
     ---------------------------------------------------------


.
. mean g030 if visit=="Y01", over (random)

Mean estimation                     Number of obs    =     2831

     Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


     ---------------------------------------------------------
         Over |       Mean   Std. Err.    [95% Conf. Interval]
     ---------+-----------------------------------------------
     g030     |
     Lifestyle |   160.6768   .8614018     158.9877    162.3658
     Metformin |   171.1185   .9128763     169.3285    172.9085
       Placebo |   173.3129   .9767071     171.3978     175.228
     ---------------------------------------------------------


.
. mean g120 if visit=="Y01", over (random)

Mean estimation                     Number of obs    =     2856

     Lifestyle: random = Lifestyle
     Metformin: random = Metformin
       Placebo: random = Placebo


     ---------------------------------------------------------
         Over |       Mean   Std. Err.    [95% Conf. Interval]
     ---------+-----------------------------------------------
     g120     |
     Lifestyle |   140.6345   1.165559     138.3491    142.9199
     Metformin |   156.2584   1.240541     153.8259    158.6908
       Placebo |   157.6926   1.385426     154.9761    160.4092
     ---------------------------------------------------------
```