

Dataset Integrity Check for The Environmental Determinants of Diabetes in the Young (TEDDY) M160 Pitchika

Prepared by Sabrina Chen

IMS Inc.

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

May 6, 2019

Contents

1 Standard Disclaimer	2
2 Study Background	2
3 Archived Datasets	2
4 Statistical Methods	2
5 Results	3
6 Conclusions	3
7 References	3
Table A: Variables used to replicate Table 1: Characteristics of the first 100 the Environmental Determinants of Diabetes in the Young (TEDDY) children diagnosed with T1D.....	3
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	5
Table C: Variables used to replicate Table 2: Symptoms and laboratory data at onset of T1D	Error! Bookmark not defined.
Table D: Comparison of values computed in integrity check to reference article Table 2 values	Error! Bookmark not defined.
Table E: Variables used to replicate Figure 2:.....	Error! Bookmark not defined.
Table F: Comparison of values computed in integrity check to reference article Figure 2	Error! Bookmark not defined.
Attachment A: SAS Code	5

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 TEDDY study was designed to follow children with and without a family history of T1D to understand the environmental factors that contribute to the disease. Newborn children younger than 4 months were screened for high-risk HLA alleles, and those with qualifying haplotypes were eligible for follow-up. Information is collected on medical information (infections, medication, immunizations), exposure to dietary and other environmental factors, negative life events, family history, tap water, and measurements of psychological stress. Biospecimens, including blood, stool, urine, and nail clippings, are taken at baseline and follow-up study visits. The primary outcome measures include two endpoints—the first appearance of one or more islet cell autoantibodies (GADA, IAA, or IA-2A), confirmed at two consecutive visits, and development of T1D. The cohort will be followed for 15 years, or until the occurrence of one of the primary endpoints.

3 Archived Datasets

All the SAS data files, as provided by the Data Coordinating Center (DCC), are located in the TEDDY folder in the data package. For this replication, variables were taken from the “m_160_apitchika_niddk_31may2017.sas7bdat” dataset.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Anitha Pitchika et al [1] in *Obesity* (2018) 26, 1457-1466. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], **Characteristics of the study population stratified according to maternal diabetes**, Table A lists the variables that were used in the replication and Table B compares the results calculated from the archived data files to the results published in Table 1. The results of the replication are an exact match to the published results.

6 Conclusions

The NIDDK repository is confident that the TEDDY data files to be distributed are a true copy of the study data.

7 References

[1] Anitha Pitchika, Kendra Vehik, Sandra Hummel, Jill M. Norris, Ulla M. Uusitalo, Jimin Yang, Suvi M. Virtanen, Sibylle Koletzko, Carin Andrén Aronsson, Anette-G. Ziegler, Andreas Beyerlein, and the TEDDY study group. Associations of Maternal Diabetes During Pregnancy with Overweight in Offspring: Results from the Prospective TEDDY Study. *Obesity* (2018) 26, 1457-1466.

Table A: Variables used to replicate Table 1: Characteristics of the study population stratified according to maternal diabetes.

Table Variable	dataset.variable
Diabetes	m_160_apitchika_niddk_31may2017.diabetes
Sex	m_160_apitchika_niddk_31may2017.gender
Country	m_160_apitchika_niddk_31may2017.country
Maternal smoking during pregnancy	m_160_apitchika_niddk_31may2017.smoker
Maternal alcohol drinking during pregnancy	m_160_apitchika_niddk_31may2017.drinker
Maternal education	m_160_apitchika_niddk_31may2017.mom_education
Gestational weight gain	m_160_apitchika_niddk_31may2017.gwg_iom
Breastfed >= 6 months	m_160_apitchika_niddk_31may2017.brstfed_6mo
Overweight at age 5.5	m_160_apitchika_niddk_31may2017.overweight
Obesity at age 5.5	m_160_apitchika_niddk_31may2017.obese
Maternal prepregnancy BMI (kg/m ²)	m_160_apitchika_niddk_31may2017.bmi_before_pregnancy
Maternal age at delivery (y)	m_160_apitchika_niddk_31may2017.maternal_age
Gestational age (wk)	m_160_apitchika_niddk_31may2017.gestational_age
Birth weight z score	m_160_apitchika_niddk_31may2017.z_birthweight
BMI SDS at age 5.5	m_160_apitchika_niddk_31may2017.z_bmi
Height SDS at age 5.5	m_160_apitchika_niddk_31may2017.z_height
Weight SDS at age 5.5	m_160_apitchika_niddk_31may2017.z_weight
Mean energy intake (kcal/d) at age 5	m_160_apitchika_niddk_31may2017.mean_energy
MVPA (min/d) at age 5	m_160_apitchika_niddk_31may2017.avg_mvpa

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

Table 1			
	Available n		
Variable	Manuscript	DSIC	Diff
Sex	5,324	5,324	0
Country	5,324	5,324	0
Maternal smoking during pregnancy	5,320	5,320	0
Maternal alcohol drinking during pregnancy	5,322	5,322	0
Maternal education	5,251	5,251	0
Gestational weight gain (according to Institute of Medicine guidelines)	5,241	5,241	0
Breastfed \geq 6 months	5,324	5,324	0
Overweight at age 5.5	5,277	5,277	0
Obesity at age 5.5	5,277	5,277	0
Maternal prepregnancy BMI (kg/m ²)	5,276	5,276	0
Maternal age at delivery	5,324	5,324	0

(y)			
Gestational age (wk)	5,318	5,318	0
Birth weight z score	5,186	5,186	0
BMI SDS at age 5.5	5,277	5,277	0
Height SDS at age 5.5	5,291	5,291	0
Weight SDS at age 5.5	5,304	5,304	0
Mean energy intake (kcal/d) at age 5	4,263	4,263	0
MVPA (min/d) at age 5	3,276	3,276	0

TOTAL							
	Category	N			Percent		
Variable		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Sex	Males	2,746	2,746	0	51.6	51.6	0
Country	US	2,013	2,013	0	37.8	37.8	0
	Finland	1,237	1,237	0	23.2	23.2	0
	Germany	274	274	0	5.2	5.1	0
	Sweden	1,800	1,800	0	33.8	33.8	0
Maternal smoking during pregnancy	Yes	497	497	0	9.3	9.3	0
Maternal alcohol drinking during pregnancy	Yes	1,831	1,831	0	34.4	34.4	0
Maternal education	High school	4,371	4,371	0	83.2	83.2	0
Gestational weight gain (according to Institute of Medicine guidelines)	Inadequate	909	909	0	17.3	17.3	0
	Adequate	1,899	1,899	0	36.2	36.2	0
	Excessive	2,433	2,433	0	46.4	46.4	0
Breastfed \geq 6 months	Yes	3,469	3,469	0	65.2	65.2	0

Overweight at age 5.5	Yes	1,154	1,154	0	21.9	21.9	0
Obesity at age 5.5	Yes	303	303	0	5.7	5.7	0
		Mean			SD		
		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Maternal prepregnancy BMI (kg/m2)		24.8	24.8	0	5.2	5.2	0
Maternal age at delivery (y)		31	31	0	4.9	4.9	0
Gestational age (wk)		39.5	39.5	0	1.6	1.6	0
Birth weight z score		0	0	0	1	1	0
BMI SDS at age 5.5		0.4	0.3	0	1	1	0
Height SDS at age 5.5		0.4	0.4	0	1	1	0
Weight SDS at age 5.5		0.5	0.5	0	1	1	0
Mean energy intake (kcal/d) at age 5		1,442.70	1442.7	0	362.2	362.2	0
MVPA (min/d) at age 5		68	68	0	34.4	34.4	0

O-nonDM							
	Category	N			Percent		
Variable		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Sex	Males	2,457	2457	0	51.6	51.6	0
Country	US	1,807	1807	0	38	38	0
	Finland	1,052	1052	0	22.1	22.1	0
	Germany	192	192	0	4	4	0
	Sweden	1,708	1708	0	35.9	35.9	0
Maternal smoking during pregnancy	Yes	426	426	0	9	9	0
Maternal alcohol drinking during pregnancy	Yes	1,623	1623	0	34.1	34.1	0
Maternal education	High school	3,877	3877	0	82.6	82.6	0
Gestational weight gain (according to Institute of Medicine guidelines)	Inadequate	754	754	0	16.1	16.1	0
	Adequate	1,725	1725	0	36.8	36.8	0
	Excessive	2,205	2205	0	47.1	47.1	0
Breastfed \geq 6 months	Yes	3,150	3150	0	66.2	66.2	0

Overweight at age 5.5	Yes	998	998	0	21.2	21.2	0
Obesity at age 5.5	Yes	252	252	0	5.3	5.3	0
		Mean			SD		
		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Maternal prepregnancy BMI (kg/m ²)		24.5	24.5	0	5	5	0
Maternal age at delivery (y)		30.9	30.9	0	4.9	4.9	0
Gestational age (wk)		39.6	39.6	0	1.6	1.6	0
Birth weight z score		-0.1	-0.1	0	1	1	0
BMI SDS at age 5.5		0.3	0.3	0	1	1	0
Height SDS at age 5.5		0.4	0.4	0	1	1	0
Weight SDS at age 5.5		0.5	0.5	0	1	1	0
Mean energy intake (kcal/d) at age 5		1,443.60	1443.6	0	359.7	359.7	0

O-GDM							
	Category	N			Percent		
Variable		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Sex	Males	174	174	0	53.4	53.4	0
Country	US	117	117	0	35.9	35.9	0
	Finland	138	138	0	42.3	42.3	0
	Germany	17	17	0	5.2	5.2	0
	Sweden	54	54	0	16.6	16.6	0
Maternal smoking during pregnancy	Yes	40	40	0	12.3	12.3	0
Maternal alcohol drinking during pregnancy	Yes	109	109	0	33.4	33.4	0
Maternal education	High school	288	288	0	89.2	89.2	0
Gestational weight gain (according to Institute of Medicine guidelines)	Inadequate	117	117	0	36.7	36.7	0
	Adequate	92	92	0	28.8	28.8	0
	Excessive	110	110	0	34.5	34.5	0
Breastfed \geq 6 months	Yes	193	193	0	59.2	59.2	0

Overweight at age 5.5	Yes	89	89	0	27.6	27.6	0
Obesity at age 5.5	Yes	32	32	0	9.9	9.9	0
		Mean			SD		
		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Maternal prepregnancy BMI (kg/m ²)		28.3	28.3	0	6.4	6.4	0
Maternal age at delivery (y)		32.2	32.2	0	5.3	5.3	0
Gestational age (wk)		39.2	39.2	0	1.7	1.7	0
Birth weight z score		0.1	0.1	0	1.1	1.1	0
BMI SDS at age 5.5		0.5	0.5	0	1.1	1.1	0
Height SDS at age 5.5		0.3	0.3	0	1	1	0
Weight SDS at age 5.5		0.5	0.5	0	1	1	0
Mean energy intake (kcal/d) at age 5		1,461.70	1461.7	0	428.5	428.5	0
MVPA (min/d) at age 5		63	63	0	33.8	33.8	0

O-T1DM							
	Category	N			Percent		
Variable		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Sex	Males	109	109	0	48.4	48.4	0
Country	US	78	78	0	34.7	34.7	0
	Finland	47	47	0	20.9	20.9	0
	Germany	65	65	0	28.9	28.9	0
	Sweden	35	35	0	15.6	15.6	0
Maternal smoking during pregnancy	Yes	30	30	0	13.3	13.3	0
Maternal alcohol drinking during pregnancy	Yes	95	95	0	42.2	42.2	0
Maternal education	High school	194	194	0	87.4	87.4	0
Gestational weight gain (according to Institute of Medicine guidelines)	Inadequate	34	34	0	15.2	15.2	0
	Adequate	78	78	0	34.9	34.8	0
	Excessive	112	112	0	50	50	0
Breastfed \geq 6 months	Yes	121	121	0	53.8	53.8	0

Overweight at age 5.5	Yes	58	58	0	25.9	25.9	0
Obesity at age 5.5	Yes	17	17	0	7.6	7.6	0
		Mean			SD		
		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Maternal prepregnancy BMI (kg/m ²)		25.3	25.3	0	4.7	4.7	0
Maternal age at delivery (y)		30.8	30.8	0	4.9	4.9	0
Gestational age (wk)		37.7	37.7	0	1.8	1.8	0
Birth weight z score		0.9	0.9	0	1.3	1.3	0
BMI SDS at age 5.5		0.4	0.4	0	1.1	1.1	0
Height SDS at age 5.5		0.3	0.3	0	0.9	0.9	0
Weight SDS at age 5.5		0.5	0.5	0	0.9	1	0
Mean energy intake (kcal/d) at age 5		1,402.10	1402.1	0	316.4	316.4	0
MVPA (min/d) at age 5		69.4	69.4	0	36.8	36.8	0

O-T2DM							
	Category	N			Percent		
Variable		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Sex	Males	6	6	0	42.9	42.9	0
Country	US	11	11	0	78.6	78.6	0
	Finland	0	0	0	0	0	0
	Germany	0	0	0	0	0	0
	Sweden	3	3	0	21.4	21.4	0
Maternal smoking during pregnancy	Yes	1	1	0	7.1	7.1	0
Maternal alcohol drinking during pregnancy	Yes	4	4	0	28.6	28.6	0
Maternal education	High school	12	12	0	85.7	85.7	0
Gestational weight gain (according to Institute of Medicine guidelines)	Inadequate	4	4	0	28.6	28.6	0
	Adequate	4	4	0	28.6	28.6	0
	Excessive	6	6	0	42.9	42.9	0
Breastfed \geq 6 months	Yes	5	5	0	35.7	35.7	0

Overweight at age 5.5	Yes	9	9	0	64.3	64.3	0
Obesity at age 5.5	Yes	2	2	0	14.3	14.3	0
		Mean			SD		
		Manuscript	DSIC	Diff	Manuscript	DSIC	Diff
Maternal prepregnancy BMI (kg/m ²)		35	35	0	7.5	7.5	0
Maternal age at delivery (y)		34	34	0	5.6	5.6	0
Gestational age (wk)		38.1	38.1	0	2.1	2.1	0
Birth weight z score		0.2	0.2	0	1	1	0
BMI SDS at age 5.5		1.1	1.1	0	1.3	1.3	0
Height SDS at age 5.5		0.2	0.2	0	0.7	0.7	0
Weight SDS at age 5.5		0.9	0.9	0	1.2	1.2	0
Mean energy intake (kcal/d) at age 5		1,353.20	1353.2	0	317.3	317.3	0
MVPA (min/d) at age 5		54.4	54.4	0	29.8	29.8	0

Attachment A: SAS Code

```
options nocenter validvarname=upcase;

title 'prj/niddk/ims_analysis/TEDDY/prog_initial_analysis/m_160_dsic.sas';
run;

*****;
* INPUT      ;
*****;

libname sasfile '/prj/niddk/ims_analysis/TEDDY/private_orig_data/M_160_APitchika_NIDDK_Submission/';

*****;
* MACROS     ;
*****;
%macro readin(ds);
  data &ds;
    set sasfile.&ds;
  run;

  proc contents data=&ds;
    title3 "&ds";
  run;
%mend;

* produce n and %;
%macro npercent(rownum, var, varf, subset, subsetname);
  proc freq data=analy noprint;
    where &subset = 1;
    tables &var/list missprint out=tbl1&subsetname;
  run;

  data tbl1&subsetname;
    length covar covarf $100;
    set tbl1&subsetname;
    covar = "&var";
    covarf = put(&var,&varf..);
    rownum = &rownum;
  run;

  data prnt&subsetname;
```

```

    set prnt&subsetname tbl1&subsetname;
run;

%mend;

%macro univ(rownum, var, subset, subsetname);

proc univariate data=analy outtable= univ&subsetname noprint;
  where &subset=1;
  var &var
  ;
run;

data univ&subsetname;
  length covarf $100;
  set univ&subsetname;
  covarf = "&subset";
  rownum = &rownum;
run;

data prntuniv&subsetname;
  set prntuniv&subsetname univ&subsetname;
run;

%mend;

*****;
* FORMATS ;
*****;
proc format;
  value novalue
    . = "No Value"
  other = " Value"
  ;

  value negpos
    0 = "Negative"
    1 = "Positive"
  ;

  value yesnof
    0 = "No"

```

```

1 = "Yes"
;

value zerof
. = "0"
;

value genderf
1 = "Male"
0 = "Female"
;

value momeducf
. = 'Missing'
0 = 'High school education or less'
1 = 'More than high school education'
;

value countryf
1 = '1 US'
2 = '2 Finland'
3 = '3 Germany'
4 = '4 Sweden'
;

value gestatf
-1 = "-1 Inadequate"
0 = " 0 Adequate"
1 = " 1 Excessive"
;

run;

%readin(m_160_apitchika_niddk_31may2017);

proc freq data=m_160_apitchika_niddk_31may2017;
  where diabetes ne .;
  tables diabetes gender country smoker drinker mom_education gwg_iom brstfed_6mo  overweight obese/missing list;
title3 "m_160_apitchika_niddk_31may2017";
run;

proc means;
  where diabetes ne .;

```

```

var bmi_before_pregnancy maternal_age gestational_age z_birthweight z_bmi z_height z_weight mean_energy avg_mvpa;
run;

data analy;
  set m_160_apitchika_niddk_31may2017 (where=(diabetes ne .));
  * create subset flag for each row to use in macro call;
  all = 1;

  if diabetes = 4 then o_nondm=1;
  else if diabetes = 1 then o_gdm=1;
  else if diabetes = 2 then o_t1dm=1;
  else if diabetes = 3 then o_t2dm=1;

run;

proc freq data=analy;
  tables diabetes*o_nondm* o_gdm* o_t1dm* o_t2dm /list missing;
run;

* n and percent;
data prntall;
  set _null_;
run;

%npercent(1, gender      , genderf      , all, all);
%npercent(2, country    , countryf  , all, all);
%npercent(3, smoker     , yesnof    , all, all);
%npercent(4, drinker   , yesnof    , all, all);
%npercent(5, mom_education, momeducf  , all, all);
%npercent(6, gwg_iom    , gestatf   , all, all);
%npercent(7, brstfed_6mo , yesnof    , all, all);
%npercent(8, overweight , yesnof    , all, all);
%npercent(9, obese      , yesnof    , all, all);

proc sort data=prntall;
  by rownum;
run;

data prntall;
  set prntall;
  by rownum;

  retain total;

```

```

if first.rownum then total=0;

if percent ne . then total = sum(total, count);
run;

proc print data=prntall;
var rownum covarf total COUNT PERCENT;
title3 "All";
run;

data prnto_nondm;
set _null_;
run;

%npcent(1, gender , genderf , o_nondm, o_nondm);
%npcent(2, country , countryf , o_nondm, o_nondm);
%npcent(3, smoker , yesnof , o_nondm, o_nondm);
%npcent(4, drinker , yesnof , o_nondm, o_nondm);
%npcent(5, mom_education, momeducf , o_nondm, o_nondm);
%npcent(6, gwg_iom , gestatf , o_nondm, o_nondm);
%npcent(7, brstfed_6mo , yesnof , o_nondm, o_nondm);
%npcent(8, overweight , yesnof , o_nondm, o_nondm);
%npcent(9, obese , yesnof , o_nondm, o_nondm);

proc print data=prnto_nondm;
var rownum covarf COUNT PERCENT;
title3 "o_nondm";
run;

data prnto_gdm;
set _null_;
run;

%npcent(1, gender , genderf , o_gdm, o_gdm);
%npcent(2, country , countryf , o_gdm, o_gdm);
%npcent(3, smoker , yesnof , o_gdm, o_gdm);
%npcent(4, drinker , yesnof , o_gdm, o_gdm);
%npcent(5, mom_education, momeducf , o_gdm, o_gdm);
%npcent(6, gwg_iom , gestatf , o_gdm, o_gdm);
%npcent(7, brstfed_6mo , yesnof , o_gdm, o_gdm);
%npcent(8, overweight , yesnof , o_gdm, o_gdm);
%npcent(9, obese , yesnof , o_gdm, o_gdm);

```

```

proc print data=prnto_gdm;
  var rownum covarf COUNT PERCENT;
title3 "o_gdm";
run;

data prnto_t1dm;
  set _null_;
run;

%npercent(1, gender      , genderf      , o_t1dm, o_t1dm);
%npercent(2, country    , countryf  , o_t1dm, o_t1dm);
%npercent(3, smoker     , yesnof    , o_t1dm, o_t1dm);
%npercent(4, drinker    , yesnof    , o_t1dm, o_t1dm);
%npercent(5, mom_education, momeducf  , o_t1dm, o_t1dm);
%npercent(6, gwg_iom    , gestatf   , o_t1dm, o_t1dm);
%npercent(7, brstfed_6mo , yesnof    , o_t1dm, o_t1dm);
%npercent(8, overweight , yesnof    , o_t1dm, o_t1dm);
%npercent(9, obese      , yesnof    , o_t1dm, o_t1dm);

proc print data=prnto_t1dm;
  var rownum covarf COUNT PERCENT;
title3 "o_t1dm";
run;

data prnto_t2dm;
  set _null_;
run;

%npercent(1, gender      , genderf      , o_t2dm, o_t2dm);
%npercent(2, country    , countryf  , o_t2dm, o_t2dm);
%npercent(3, smoker     , yesnof    , o_t2dm, o_t2dm);
%npercent(4, drinker    , yesnof    , o_t2dm, o_t2dm);
%npercent(5, mom_education, momeducf  , o_t2dm, o_t2dm);
%npercent(6, gwg_iom    , gestatf   , o_t2dm, o_t2dm);
%npercent(7, brstfed_6mo , yesnof    , o_t2dm, o_t2dm);
%npercent(8, overweight , yesnof    , o_t2dm, o_t2dm);
%npercent(9, obese      , yesnof    , o_t2dm, o_t2dm);

proc print data=prnto_t2dm;
  var rownum covarf COUNT PERCENT;
title3 "o_t2dm";
run;

```

```

* combine rows;
proc sort data=prntall;
  by rownum covarf;
run;

proc sort data=prnto_nondm (rename=(count   = o_nondm_count
                                   percent = o_nondm_percent));
  by rownum covarf;
run;

proc sort data=prnto_gdm (rename=(count   = o_gdm_count
                                   percent = o_gdm_percent));
  by rownum covarf;
run;

proc sort data=prnto_t1dm (rename=(count   = o_t1dm_count
                                   percent = o_t1dm_percent));
  by rownum covarf;
run;

proc sort data=prnto_t2dm (rename=(count   = o_t2dm_count
                                   percent = o_t2dm_percent));
  by rownum covarf;
run;

data allnprcnt;
  merge prntall      (in=in1 keep=rownum covar covarf count percent total)
        prnto_nondm (in=in2 keep=rownum covar covarf o_nondm_count o_nondm_percent)
        prnto_gdm   (in=in3 keep=rownum covar covarf o_gdm_count o_gdm_percent)
        prnto_t1dm  (in=in4 keep=rownum covar covarf o_t1dm_count o_t1dm_percent)
        prnto_t2dm  (in=in5 keep=rownum covar covarf o_t2dm_count o_t2dm_percent)
        ;
  by rownum covarf;
  if in1 or in2 or in3 or in4 or in5;
run;

* med, min and max;

data prntunivall;
  set _null_;

```

```

run;

%univ(10, bmi_before_pregnancy , all, all);
%univ(11, maternal_age , all, all);
%univ(12, gestational_age , all, all);
%univ(13, z_birthweight , all, all);
%univ(14, z_bmi , all, all);
%univ(15, z_height , all, all);
%univ(16, z_weight , all, all);
%univ(17, mean_energy , all, all);
%univ(18, avg_mvpa , all, all);

proc print data= prntunivall noobs;
  var rownum _var_ covarf _nobs_ _median_ _min_ _max_ _std_;
run;

data prntunivo_nondm;
  set _null_;
run;

%univ(10, bmi_before_pregnancy , o_nondm, o_nondm);
%univ(11, maternal_age , o_nondm, o_nondm);
%univ(12, gestational_age , o_nondm, o_nondm);
%univ(13, z_birthweight , o_nondm, o_nondm);
%univ(14, z_bmi , o_nondm, o_nondm);
%univ(15, z_height , o_nondm, o_nondm);
%univ(16, z_weight , o_nondm, o_nondm);
%univ(17, mean_energy , o_nondm, o_nondm);
%univ(18, avg_mvpa , o_nondm, o_nondm);

proc print data= prntunivo_nondm noobs;
  var rownum _var_ covarf _nobs_ _mean_ _min_ _max_ _std_;
run;

data prntunivo_gdm;
  set _null_;
run;

%univ(10, bmi_before_pregnancy , o_gdm, o_gdm);
%univ(11, maternal_age , o_gdm, o_gdm);
%univ(12, gestational_age , o_gdm, o_gdm);
%univ(13, z_birthweight , o_gdm, o_gdm);
%univ(14, z_bmi , o_gdm, o_gdm);

```

```

%univ(15, z_height          , o_gdm, o_gdm);
%univ(16, z_weight         , o_gdm, o_gdm);
%univ(17, mean_energy      , o_gdm, o_gdm);
%univ(18, avg_mvpa         , o_gdm, o_gdm);

proc print data= prntunivo_gdm noobs;
  var rownum _var_ covarf _nobs_ _mean_ _min_ _max_ _std_;
run;

data prntunivo_t1dm;
  set _null_;
run;

%univ(10, bmi_before_pregnancy , o_t1dm, o_t1dm);
%univ(11, maternal_age         , o_t1dm, o_t1dm);
%univ(12, gestational_age      , o_t1dm, o_t1dm);
%univ(13, z_birthweight        , o_t1dm, o_t1dm);
%univ(14, z_bmi                , o_t1dm, o_t1dm);
%univ(15, z_height             , o_t1dm, o_t1dm);
%univ(16, z_weight             , o_t1dm, o_t1dm);
%univ(17, mean_energy          , o_t1dm, o_t1dm);
%univ(18, avg_mvpa             , o_t1dm, o_t1dm);

proc print data= prntunivo_t1dm noobs;
  var rownum _var_ covarf _nobs_ _mean_ _min_ _max_ _std_;
run;

data prntunivo_t2dm;
  set _null_;
run;

%univ(10, bmi_before_pregnancy , o_t2dm, o_t2dm);
%univ(11, maternal_age         , o_t2dm, o_t2dm);
%univ(12, gestational_age      , o_t2dm, o_t2dm);
%univ(13, z_birthweight        , o_t2dm, o_t2dm);
%univ(14, z_bmi                , o_t2dm, o_t2dm);
%univ(15, z_height             , o_t2dm, o_t2dm);
%univ(16, z_weight             , o_t2dm, o_t2dm);
%univ(17, mean_energy          , o_t2dm, o_t2dm);
%univ(18, avg_mvpa             , o_t2dm, o_t2dm);

proc print data= prntunivo_t2dm noobs;
  var rownum _var_ covarf _nobs_ _mean_ _min_ _max_ _std_;

```

```

run;

* combine rows;
proc sort data=prntunivall (rename=( _nobs_=total));
  by rownum covarf;
run;

proc sort data=prntunivo_nondm (rename=( _mean_ = o_nondm_mean_
                                         _std_   = o_nondm_std_ ))
  ;
  by rownum covarf;
run;

proc sort data=prntunivo_gdm (rename=( _mean_ = o_gdm_mean_
                                         _std_   = o_gdm_std_ ))
  ;
  by rownum covarf;
run;

proc sort data=prntunivo_t1dm (rename=( _mean_ = o_t1dm_mean_
                                         _std_   = o_t1dm_std_ ))
  ;
  by rownum covarf;
run;

proc sort data=prntunivo_t2dm (rename=( _mean_ = o_t2dm_mean_
                                         _std_   = o_t2dm_std_ ))
  ;
  by rownum covarf;
run;

data alluniv;
  merge prntunivall      (in=in1 keep = rownum _var_ covarf _mean_ _std_ total)
        prntunivo_nondm (in=in2 keep = rownum _var_ covarf o_nondm_mean_
                          o_nondm_std_)
        prntunivo_gdm   (in=in3 keep = rownum _var_ covarf o_gdm_mean_
                          o_gdm_std_)
        prntunivo_t1dm  (in=in4 keep = rownum _var_ covarf o_t1dm_mean_
                          o_t1dm_std_)
        prntunivo_t2dm  (in=in5 keep = rownum _var_ covarf o_t2dm_mean_
                          o_t2dm_std_)
  ;
  by rownum;

```

```

    if in1 or in2;
run;

* set all rows together and format;
data table;
  set allnprcnt
      alluniv;

      percent          = round(percent          , 0.1);
  o_nondm_percent     = round(o_nondm_percent, 0.1);
  o_gdm_percent       = round(o_gdm_percent   , 0.1);
  o_t1dm_percent      = round(o_t1dm_percent  , 0.1);
  o_t2dm_percent      = round(o_t2dm_percent  , 0.1);
  _mean_              = round(_mean_         , 0.1);
  o_nondm_mean_       = round(o_nondm_mean_   , 0.1);
  o_gdm_mean_         = round(o_gdm_mean_     , 0.1);
  o_t1dm_mean_        = round(o_t1dm_mean_    , 0.1);
  o_t2dm_mean_        = round(o_t2dm_mean_    , 0.1);
  _std_               = round(_std_          , 0.1);
  o_nondm_std_        = round(o_nondm_std_    , 0.1);
  o_gdm_std_          = round(o_gdm_std_      , 0.1);
  o_t1dm_std_         = round(o_t1dm_std_     , 0.1);
  o_t2dm_std_         = round(o_t2dm_std_     , 0.1);

run;

proc sort data=table;
  by rownum covarf;
run;

proc print data=table;
  where (1 <= rownum <= 9) and covarf ne "No";
  var rownum _var_ covar covarf total
      count          percent
      o_nondm_count o_nondm_percent
      o_gdm_count   o_gdm_percent
      o_t1dm_count  o_t1dm_percent
      o_t2dm_count  o_t2dm_percent;
  title3 "Table 1 - n(%)";
run;

proc print data=table;
  where (10 <= rownum <= 18);
  var rownum _var_ covar covarf total

```

```
      _mean_      _std_  
o_nondm_mean_  o_nondm_std_  
o_gdm_mean_    o_gdm_std_  
o_t1dm_mean_   o_t1dm_std_  
o_t2dm_mean_   o_t2dm_std_  
title3 "Table 1 - mean+-sd";  
run;
```