

# Dataset Integrity Check for Effect of Oral Insulin on Prevention of Diabetes in Relatives of Patients With Type 1 Diabetes Krischer

Prepared by Jane Rideau Demuth

IMS Inc.

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

August 30, 2018

## 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 Figure 1: Participant Flow Through Type 1 Diabetes TrialNet Study ..	4
Table B: Comparison of values computed in integrity check to reference article Figure 1 values.....	5
Table C: Variables used to replicate Table 1: Distribution of Participant Characteristics by Treatment Group and Strata	
Table D: Comparison of values computed in integrity check to reference article Table 1 values	
Attachment A: SAS Code .....	6

## **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**

In the Diabetes Prevention Trial–Type 1 (DPT-1), oral insulin compared with placebo did not show a reduction in the development of diabetes, but a post hoc analysis identified an at-risk subgroup with higher insulin autoantibody titers that suggested benefit. Consequently, the Type 1 Diabetes TrialNet clinical trials network, which succeeded the DPT-1 trial group, sought to further explore the role of oral insulin in delaying diabetes among relatives who were not significantly different from those in the subgroup who had experienced apparent benefit from oral insulin in DPT-1.

## **3 Archived Datasets**

All the SAS data files, as provided by the Data Coordinating Center (DCC), are located in the TrialNet\_07/private\_orig\_data/Data.Extraction.7/7/sasv9/ folder in the data package. For this replication, variables were taken from the sas7bdat datasets ae, autoab, bmi, cpepm, cpepo, gtt hba1c, hla, insulin, mastable, rel, and rx.

## **4 Statistical Methods**

Analyses were performed to duplicate results for the data published by Kirscher et al [1] in the journal JAMA Care in 2017. To verify the integrity of the dataset, descriptive statistics were computed.

## 5 Results

For Figure 1 in the publication [1], Participant Flow Through Type 1 Diabetes TrialNet Study, 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 Figure 1.

For Table 1 in the publication [1], Distribution of Participant Characteristics by Treatment Group and Strata, Table C lists the variables that were used in the replication and Table D compares the results calculated from the archived data files to the results published in Table 1.

## 6 Conclusions

The NIDDK repository is confident that the TrialNet\_07 data files to be distributed are a true copy of the study data.

## 7 References

[1] Writing Committee for the Type 1 Diabetes TrialNet Oral Insulin Study Group. Effect of Oral Insulin on Prevention of Diabetes in Relatives of Patients With Type 1 DiabetesA Randomized Clinical Trial. *JAMA*. 2017;318(19):1891–1902. doi:10.1001/jama.2017.17070.

**Table A:** Variables used to replicate Figure 1: Participant Flow Through Type 1 Diabetes TrialNet Study

<b>Table Variable</b>	<b>dataset.variable</b>
Randomized	mastable.reg
Randomized to receive	mastable.rxdesc
Stratum	mastable.stratum2

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

REG value	RXDESC value	STRATUM2 value	MS n	IMS n	Diff n (%)
Randomized	-	-	560	560	0 (0)
	Oral Insulin	-	283	283	0 (0)
		Primary Stratum	203	202	1 (0.49)
		Secondary Stratum 1	28	30	2 (7.4)
		Secondary Stratum 2	51	51	0 (0)
		Secondary Stratum 3	1	0	1 (100)
	Placebo	-	277	277	0 (0)
		Primary Stratum	186	190	4 (2.2)
		Secondary Stratum 1	27	26	1 (3.7)
		Secondary Stratum 2	63	59	4 (6.3)
		Secondary Stratum 3	1	2	1 (100)

**Table C:** Variables used to replicate Table 1: Distribution of Participant Characteristics by Treatment Group and Strata

<b>Table Variable</b>	<b>dataset.variable</b>
Randomized	mastable.reg
Randomized to receive	mastable.rxdesc
Stratum	mastable.stratum2
Sex	mastable.sex
Race	mastable.race
Ethnicity	mastable.ethnicity
Height	bmi.heightcm
Weight	bmi.weightkg

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

<b>STRATUM2 value (cohort)</b>	<b>RXDESC value (active/placebo)</b>	<b>Participant Characteristic</b>	<b>MS n (%) or mdn (IQR)</b>	<b>IMS n (%) or mdn (IQR)</b>	<b>Diff n (%)</b>		
Entire cohort	Oral insulin (active)	-	283	283	0		
		Age (median (IQR))	8.2 (5.9-12.5)	8.2 (5.9-12.6)	0 (0-0.1)		
		Gender = Boys	170 (60.1)	170 (60.1)	0 (0)		
		Race = White	252 (95.5)	252 (95.5)	0 (0)		
		Race = Black	8 (3.0)	8 (3.0)	0 (0)		
		Race = Asian/Pacific Islander	4 (1.5)	4 (1.5)	0 (0)		
		Ethnicity = Non-Hispanic	256 (90.5)	256 (90.5)	0 (0)		
		Placebo	Placebo	-	277	277	0
				Age (median (IQR))	8.2 (5.4-11.5)	8.2 (5.4-11.5)	0 (0-0)
				Gender = Boys	170 (61.4)	170 (61.4)	0 (0)
Race = White	249 (94.3)			250 (94.3)	1 (0)		
Race = Black	9 (3.4)			9 (3.4)	0 (0)		
Race = Asian/Pacific Islander	6 (2.3)			6 (2.3)	0 (0)		
Ethnicity = Non-Hispanic	252 (91.0)			253 (91.3)	1 (0.3)		
Primary Stratum	Oral insulin (active)			-	203	202	1
				Age (median (IQR))	8.6 (6.1-12.8)	8.6 (6.1-12.7)	0 (0-0.1)
				Gender = Boys	128 (63.1)	126 (62.4)	2 (0.7)
		Race = White	181 (95.3)	180 (95.2)	1 (0.1)		
		Race = Black	6 (3.2)	6 (3.2)	0 (0)		
		Race = Asian/Pacific Islander	3 (1.6)	3 (1.6)	0 (0)		
		Ethnicity = Non-Hispanic	182 (89.7)	181 (89.6)	1 (0.1)		
		Placebo	Placebo	-	186	190	4
				Age (median (IQR))	8.2 (5.5-11.8)	8.2 (5.5-11.5)	0 (0-0.3)
				Gender = Boys	117 (62.9)	119 (62.6)	2 (0.3)



<b>STRATUM2 value (cohort)</b>	<b>RXDESC value (active/placebo)</b>	<b>Participant Characteristic</b>	<b>MS n (%) or mdn (IQR)</b>	<b>IMS n (%) or mdn (IQR)</b>	<b>Diff n (%)</b>
		Race = White	172 (94.5)	176 (94.6)	4 (0.1)
		Race = Black	4 (2.2)	4 (2.2)	0 (0)
		Race = Asian/Pacific Islander	6 (3.3)	6 (3.2)	0 (0.1)
		Ethnicity = Non-Hispanic	171 (91.9)	174 (91.6)	3 (0.3)
Secondary Stratum 1	Oral insulin (active)	-	28	30	2
		Age (median (IQR))	9.1 (5.9-13.7)	8.9 (6.1-14.2)	0.2 (0.2-0.5)
		Gender = Boys	19 (67.9)	21 (70.0)	2 (2.1)
		Race = White	25 (96.2)	27 (96.4)	2 (0.2)
		Race = Black	0	0	0
		Race = Asian/Pacific Islander	1 (3.8)	1 (3.7)	0 (0.1)
		Ethnicity = Non-Hispanic	26 (92.9)	28 (93.3)	2 (0.4)
	Placebo	-	27	26	1
		Age (median (IQR))	8.5 (6.5-10.8)	8.5 (6.7-10.7)	0 (0.2-0.1)
		Gender = Boys	19 (70.4)	18 (69.2)	1 (1.2)
		Race = White	25 (100)	24 (100)	1 (0)
		Race = Black	0	0	0
		Race = Asian/Pacific Islander	0	0	0
		Ethnicity = Non-Hispanic	26 (96.3)	25 (96.2)	1 (0.1)
Secondary Stratum 2 & 3	Oral insulin (active)	-	52	51	1
		Age (median (IQR))	7.3 (5.1-10.3)	7.2 (4.7-10.2)	0.1 (0.4-0.1)
		Gender = Boys	23 (44.2)	23 (45.1)	0 (0.9)
		Race = White	46 (95.8)	45 (95.7)	1 (0.1)
		Race = Black	2 (4.2)	2 (4.3)	0 (0.1)
		Race = Asian/Pacific Islander	0	0	0
		Ethnicity = Non-Hispanic	48 (92.3)	47 (92.2)	1 (0.1)
	Placebo	-	64	61	3

<b>STRATUM2 value (cohort)</b>	<b>RXDESC value (active/placebo)</b>	<b>Participant Characteristic</b>	<b>MS n (%) or mdn (IQR)</b>	<b>IMS n (%) or mdn (IQR)</b>	<b>Diff n (%)</b>
		Age (median (IQR))	8.3 (5.1-11.5)	8.3 (5.4-11.7)	0 (0.3-0.2)
		Gender = Boys	34 (53.1)	33 (54.1)	1 (1.0)
		Race = White	52 (91.2)	50 (90.9)	2 (0.3)
		Race = Black	5 (8.8)	5 (9.1)	0 (0.3)
		Race = Asian/Pacific Islander	0	0	0
		Ethnicity = Non-Hispanic	55 (85.9)	54 (88.5)	1 (2.6)

## Attachment A: SAS Code

```
options mprint nocentre linesize=163 validvarname=upcase;

%let rundate = y2018m09d06;
%let olddate = yYYYYmMMdDD;

title "Program: /prj/niddk/ims_analysis/TrialNet_07/prog_initial_analysis/Data.Extraction.7/7/DSIC.paper.review.&rundate..sas";
title2 "This program reviews TrialNet_07/Data.Extraction.7/7 data for level 1 PII";

/*****

programmer: Jane Rideau Demuth

platform: LINUX SASv9.4

date: 06 September 2018

purpose: See title2.

*****/

*****;
*** formats ***;
*****;
proc format;
  value nmsgf
    . = ' '
    low-high = '###'
  ;
  value $cmsgf
    ' ' = ' '
    other = '$$$'
  ;
  value $stratumf
    'Primary' = 'Primary'
    'Secondary 1' = 'Secondary 1'
    'Secondary 2','Secondary 3' = 'Secondary 2&3'
  ;

*****;
*** input files ***;
*****;
libname tn777 "/prj/niddk/ims_analysis/TrialNet_07/private_orig_data/Data.Extraction.7/7/sasv9/";
data mastable;
  set tn777.mastable;
title3 "Input file: /prj/niddk/ims_analysis/TrialNet_07/private_orig_data/Data.Extraction.7/7/sasv9/mastable.sas7bdat";
proc contents data=mastable varnum;
```

```

%macro readin(ds);
data &ds.;
  set tn777.&ds.;
title3 "Input file: /prj/niddk/ims_analysis/TrialNet_07/private_orig_data/Data.Extraction.7/7/sasv9/";
title4 "          ~&ds..sas7bdat";
proc contents data=&ds. varnum;
%mend readin;
%readin(ae);
%readin(autoab);
%readin(bmi);
%readin(cpepm);
%readin(cpepo);
%readin(gtt);
%readin(hbalc);
%readin(hla);
%readin(insulin);
%readin(rel);
%readin(rx);

*****;
*** check figure 1 ***;
*****;
proc freq data=mastable;
  title3 'Figure 1';
  tables reg*rxdesc*stratum2 / missing list;

*****;
*** check table ***;
*****;
data mastable;
  set mastable;

  *** calculate age ***;
  age = (dtreg-dtbirth)/365.25;

proc sort data=mastable;
  by stratum2;

proc freq data=mastable;
  title3 'Table';
  where reg = 0;
  tables rxdesc*(sex race ethnic) / missing;

proc freq data=mastable;
  title3 'Table';
  where reg = 0;
  by stratum2;
  tables rxdesc*(sex race ethnic) / missing;
  format stratum2 $stratumf.;

proc sort data=bmi;

```

```

by maskid dtexam;

data bmlst(keep=maskid bmi);
  set bmi(keep=maskid dtexam weightkg heightcm);
  by maskid dtexam;
  if first.maskid;

  *** calculate BMI ***;
  bmi = weightkg / ((heightcm/100) ** 2);

proc sort data=mastable;
  by maskid;

data mastable;
  merge mastable(in=keepthese)
        bmlst(in=inbmi keep=maskid bmi);
  by maskid;
  if not(keepthese and inbmi) then abort;

proc sort data=mastable;
  by rxdesc;

proc univariate data=mastable noprint;
  var age bmi;
  by rxdesc;
  output out=mastableuni
         median=age_mdn bmi_mdn
         q1=age_q25 bmi_q25
         q3=age_q75 bmi_q75;

proc print data=mastableuni noobs;
  title3 'Table - Entire Cohort by Treatment';
  var rxdesc age_q25 age_mdn age_q75 bmi_q25 bmi_mdn bmi_q75;
  format age_q25 age_mdn age_q75 8.1;

proc sort data=mastable;
  by stratum2 rxdesc;
  format stratum2 $stratumf.;

proc univariate data=mastable noprint;
  var age bmi;
  by stratum2 rxdesc;
  output out=mastableuni2
         median=age_mdn bmi_mdn
         q1=age_q25 bmi_q25
         q3=age_q75 bmi_q75;
  format stratum2 $stratumf.;

proc print data=mastableuni2 noobs;
  title3 'Table - Entire Cohort by Stratum, Treatment';
  var stratum2 rxdesc age_q25 age_mdn age_q75 bmi_q25 bmi_mdn bmi_q75;

```

```
format age_q25 age_mdn age_q75 8.1;
```

```
endsas;
```