Dataset Integrity Check for Immune Effects of Oral Insulin in Relatives at Risk for Type 1 Diabetes Mellitus (TN20 IEOI) Study Data

Prepared by NIDDK-CR
November 16, 2022

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 TrialNet (TN20) study is a two-arm, multi-center, randomized, open-labeled clinical trial designed to assess the effects of varying doses and schedules of oral insulin on immunological and metabolic markers in relatives at risk for type 1 diabetes.

Participants are from the TrialNet Natural History/Pathway to Prevention (TN01) study and must have a relative with type 1 diabetes and be positive for insulin autoantibodies and at least one other autoantibody. All participants receive an active treatment of recombinant insulin in capsules of either 67.5 mg daily or 500 mg every other week. Participants visit the study site up to 11 times over one year for blood tests and other study procedures. During the beginning treatment phase, there are two visits one month apart for those given the 67.5 mg dose, and three visits two weeks apart to titrate the dose for those in the 500 mg treatment group. There are three additional treatment visits at months 2, 3, and 6, as well as four additional visits for follow-up at months 7, 8, 9, and 12. The primary outcome is the change in immune function as assessed by change in level or quality of T lymphocyte or autoantibody biomarkers measured between 13 and 26 weeks compared to the baseline.

3 Archived Datasets

All data files, as provided by the Data Coordinating Center (DCC), are located in the TN20 folder in the data package. For this replication, variables were taken from the “tn20_treatmentstartdate”, “tn01_nh01_screening”, “tn20_ab_rba_ecl_aff_subtyping”, “tn20_diabetesonset”, and “tn20_initialvisit” datasets.

4 Statistical Methods

Analyses were performed to replicate descriptive statistics provided by the DCC for the TN20 study. To verify the integrity of the datasets, only descriptive statistics were computed.

5 Results

For the descriptive statistics table provided by the DCC, Subject Characteristics at Baseline, by Treatment Group, 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 in the provided table. The results of the replication are within expected variation of the provided results.

6 Conclusions

The NIDDK Central Repository is confident that the TN20 data files to be distributed are a true copy of the study data.

Table A: Variables used to replicate the descriptive statistics table from the DCC – Subject Characteristics at Baseline, by Treatment Group

Table Variable	**dataset.variable**
Age	tn20_treatmentstartdate.treatmentdesc tn01_nh01_screening.age
Male sex	tn20_treatmentstartdate.treatmentdesc tn01_nh01_screening.sex
Race	tn20_treatmentstartdate.treatmentdesc tn01_nh01_screening.race_white tn01_nh01_screening.race_asian tn01_nh01_screening.race_americanindianoralaskanat tn01_nh01_screening.race_blackorafricanamerican tn01_nh01_screening.race_nativehawaiinorotherpac
Ethnicity	tn20_treatmentstartdate.treatmentdesc tn01_nh01_screening.ethnicity
Autoantibodies positive	tn20_treatmentstartdate.treatmentdesc tn20_ab_rba_ecl_aff_subtyping.gad65h tn20_ab_rba_ecl_aff_subtyping.ia_2h tn20_ab_rba_ecl_aff_subtyping.znt8 tn20_ab_rba_ecl_aff_subtyping.miaa
Glycated hemoglobin	tn20_treatmentstartdate.treatmentdesc tn20_diabetesonset.HbA1cResult
Body Mass Index (kg/m^2)	tn20_treatmentstartdate.treatmentdesc tn20_initialvisit.height tn20_initialvisit.weight

Table B: Comparison of values computed in integrity check to reference table from the DCC

Subject Characteristics	Oral Insulin (500 mg every other week) (n=47)	DSIC: Oral Insulin (500 mg every other week) (n=47)	Diff. (n=0)	Oral Insulin (67.5 mg daily) (n=45)	DSIC: Oral Insulin (67.5 mg daily) (n=45)	Diff. (n=0)
Age (years)						
Median (Q1-Q3)	7 (5-11)	6 (5-9)	1 (0-2)	8 (6-10)	6 (3-9)	2 (3-1)
Range	7 (3-43)	6 (1-43)	1 (2-0)	8 (3-28)	6 (1-26)	2 (2-2)
Male sex	26 (55.3%)	26 (55.3%)	0 (0)	25 (55.6%)	25 (55.6%)	0 (0)
Race						
White	42 (89.4%)	42 (89.4%)	0 (0)	42 (93.3%)	42 (93.3%)	0 (0)
African American	3 (6.4%)	3 (6.4%)	0 (0)	1 (2.2%)	1 (2.2%)	0 (0)
Asian	1 (2.1%)	1 (2.1%)	0 (0)	0 (0.0%)	0 (0.0%)	0 (0)
More than One Race	1 (2.1%)	1 (2.1%)	0 (0)	1 (2.2%)	1 (2.2%)	0 (0)
Unknown	0 (0.0%)	0 (0.0%)	0 (0)	1 (2.2%)	1 (2.2%)	0 (0)
Ethnicity						
Non-Hispanic	41 (87.2%)	41 (87.2%)	0 (0)	43 (95.6%)	42 (95.4%)	1 (0.2)
Autoantibodies Positive						
Anti-GAD65 (harmonized)	41 (87.2%)	39 (83.0%)	2 (4.2)	41 (91.1%)	44 (97.8%)	3 (6.7)
Micro Insulin	46 (97.9%)	38 (80.8%)	8 (17.1)	45 (100.0%)	44 (97.8%)	1 (2.2)
Anti-IA-2 (harmonized)	24 (51.1%)	25 (53.2%)	1 (2.1)	23 (51.1%)	27 (60.0%)	4 (8.9)
Zinc Transporter	29 (61.7%)	32 (68.1%)	3 (6.4)	28 (62.2%)	37 (82.2%)	9 (20.0)
Glycated hemoglobin (%)						
Median (Q1-Q3)	5.1 (4.9-5.2)	5.6 (5.4-5.8)	0.5 (0.5-0.6)	5.1 (4.8-5.2)	6.9 (6.9-6.9)	1.8 (2.1-1.7)
Body Mass Index (kg/m^2)						
Median (Q1-Q3)	16.94 (15.51-20.53)	17.08 (15.78-20.53)	0.14 (0.27-0)	16.69 (15.21-19.84)	16.69 (15.21-19.84)	0 (0-0)

Attachment A: SAS Code

```
libname tn20 "X:\NIDDK\niddk-
dr_studies6\TrialNet_20\private_orig_data\Data.Extraction.20.Version20210408PW\20\sasv9";
libname tn01 "X:\NIDDK\niddk-dr_studies6\TrialNet_01\private_created_data\TN01_V6\Data";

/**********************/
/* DSIC for TN20 */
/**********************/

*creating temp datasets;
data treat; set tn20.tn20_treatmentstartdate;
run;

data v1; set tn20.tn20_initialvisit;
run;

data screen; set tn01.tn01_nh01_screening;
run;

data ab; set tn20.tn20_ab_rba_ecl_aff_subtyping;
run;

*merging;
proc sort data=treat;
by MaskID;
run;

proc sort data=screen;
by MaskID;
run;

data one; merge
treat (in=a)
screen (in=b);
by maskid;
if a=1;
run;

*total participants in each treatment arm;
proc freq data=one;
tables TreatmentDesc;
run;

*Age in years;
proc means data=one n median q1 q3 min max;
var age;
```

```
class TreatmentDesc;
run;

*sex;
proc freq data=one;
tables sex*TreatmentDesc/norow nopercent;
run;

*RACE;
data multiple; set one;
multiple_race = 0;
if race_white + race_asian + race_americanindianoralaskanat + race_blackorafricanamerican +
race_nativehawaiianorotherpaci >= 2
then multiple_race = 1;
run;

data multiple_1; set multiple;
if multiple_race = 1 AND race_white = 1 then race_white = 0;
if multiple_race = 1 AND race_blackorafricanamerican = 1 then race_blackorafricanamerican = 0;
if multiple_race = 1 AND race_americanindianoralaskanat = 1 then race_americanindianoralaskanat = 0;
if multiple_race = 1 AND race_asian = 1 then race_asian = 0;
if multiple_race = 1 AND race_americanindianoralaskanat = 1 then race_americanindianoralaskanat = 0;
run;

proc freq data=multiple_1;
tables (Race_White Race_BlackorAfricanAmerican Race_Asian Race_Unknownornotreported
multiple_race)*TreatmentDesc/norow nopercent;
run;

*ethnicity;
proc freq data=one;
tables Ethnicity*TreatmentDesc/norow nopercent;
run;

*autoantibodies positive;
proc sort data=ab;
by MaskID;
run;

proc sort data=one;
by MaskID;
run;

data two; merge
one (in=a)
ab (in=b);
by maskid;
run;
```

```
data three; set two;
if gad65h > 20 then gad = 1; else gad = 0;
if ia_2h > 5 then ia2h = 1; else ia2h = 0;
if znt8 > 0.02 then zn = 1; else zn = 0;
if miaa > 0.01 then mi = 1; else mi = 0;
run;

*GAD;
data gad; set three;
where gad = 1;
keep maskid gad;
run;

proc sort data=gad nodup;
by maskid;
run;

data gad_1; merge
treat (in=a)
gad (in=b);
by maskid;
if a=1;
run;

proc freq data=gad_1;
tables gad*TreatmentDesc/missing norow nopercent;
run;

*miaa;
data mi; set three;
where mi = 1;
keep maskid mi;
run;

proc sort data=mi nodup;
by maskid;
run;

data mi_1; merge
treat (in=a)
mi (in=b);
by maskid;
if a=1;
run;

proc freq data=mi_1;
tables mi*treatmentdesc/missing norow nopercent;
```

```
run;

*ia2h;
data ia2h; set three;
where ia2h = 1;
keep MaskID ia2h;
run;
proc sort data=ia2h nodup;
by maskid;
run;

data ia2h_1; merge
treat (in=a)
ia2h (in=b);
by maskid;
if a=1;
run;

proc freq data=ia2h_1;
tables ia2h*treatmentdesc/missing norow nopercent;
run;

*znt8;
data znt8; set three;
where zn = 1;
keep maskid zn;
run;

proc sort data=znt8 nodup;
by maskid;
run;

data znt8_1; merge
treat (in=a)
znt8 (in=b);
by maskid;
if a=1;
run ;

proc freq data=znt8_1;
tables zn*treatmentdesc/missing norow nopercent;
run;

*Hba1c;
data diab; set tn20.tn20_diabetesonset;
run;

proc sort data=diab;
```

```
by maskid;
run;

data hb; merge
treat (in=a)
diab (in=ab);
by maskid;
if a=1;
run;

proc means data=hb median q1 q3;
var HbA1cResult;
class TreatmentDesc;
run;

*BMI;
data initial; set tn20.tn20_initialvisit;
run;

proc sort data=initial;
by maskid;
run;

data bmi; merge
treat (in=a)
initial (in=b);
by maskid;
if a=1;
run;

data bmi1; set bmi;
height_m = height/100;
bmi = weight/ (height_m*height_m);
run;

proc means data=bmi1 median q1 q3;
var bmi;
class TreatmentDesc;
run;
```