Dataset Integrity Check for Anti-Thymocyte Globulin (ATG) and Pegylated Granulocyte Colony Stimulating Factor (GCSF) in New Onset Type 1 Diabetes (TrialNet19 ATG-GCSF) Study Data

Prepared by NIDDK-CR
August 19, 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 TrialNet19 (TN19) study was a three-arm, 1:1:1 randomized, placebo controlled, double-blinded trial in which participants either received active Anti-Thymocyte Globulin (ATG) and Granulocyte Colony Stimulating Factor (GCSF), received ATG alone, or received a placebo alone within 100 days from diagnosis of type 1 diabetes. TrialNet researchers assessed whether ATG used alone or in combination with GCSF helped participants continue to make some of their own insulin.

The study had a treatment phase and a follow-up phase. The treatment phase was during the first 3 months of the study. During this time, participants had one inpatient stay for 3 days and 2 nights to receive two infusions of low dose ATG or placebo followed by one injection of GCSF or placebo. Participants returned for five additional outpatient visits over the next 10 weeks (every 2 weeks) to receive an injection of GCSF or placebo. After the treatment phase, participants moved to the follow-up phase and returned for outpatient visits every 3 to 6 months. The study had a total participation time of two years.

3 Archived Datasets

All data files, as provided by the Data Coordinating Center (DCC), are located in the TN19 folder in the data package. For this replication, variables were taken from the "tn19_screeningvisitassessment.sas7bdat", "tn19_researchlabs.sas7bdat dataset", and "tn19_treatmentstartdate.sas7bdat" datasets.

4 Statistical Methods

Analyses were performed to replicate results for the data in the publication by Haller et al. [1]. To verify the integrity of the data, only descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline patient characteristics, 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 Table 1. The results of the replication are within expected variation of the published results.

6 Conclusions

The NIDDK Central Repository is confident that the TN19 data files to be distributed are a true copy of the study data.

7 References

[1] Haller MJ, Long SA, Blanchfield JL, Schatz DA, Skyler JS, Krischer JP, Bundy BN, Geyer SM, Warnock MV, Miller JL, Atkinson MA, Becker DJ, Baidal DA, DiMeglio LA, Gitelman SE, Goland R, Gottlieb PA, Herold KC, Marks JB, Moran A, Rodriguez H, Russell WE, Wilson DM, Greenbaum CJ. Low-Dose Anti-Thymocyte Globulin Preserves C-Peptide, Reduces HbA1c, and Increases Regulatory to Conventional T-Cell Ratios in New-Onset Type 1 Diabetes: Two-Year Clinical Trial Data. Diabetes, 68(6), 1267-1276, April 2019. doi: https://doi.org/10.2337/db19-0057

Table A: Variables used to replicate Table 1 – Baseline patient characteristics

Table Variable	**dataset.variable**
Age (years)	tn19_treatmentstartdate.TreatmentDesc tn19_screeningvisitassessment.DateOfBirthMonth tn19_screeningvisitassessment.DateOfBirthYear tn19_screeningvisitassessment.VisitDt
Male sex	tn19_treatmentstartdate.TreatmentDesc tn19_screeningvisitassessment.SexGenderPerson
Race	tn19_treatmentstartdate.TreatmentDesc tn19_screeningvisitassessment.Race_White tn19_screeningvisitassessment.Race_BlackOrAfricanAmerican
Ethnicity	tn19_treatmentstartdate.TreatmentDesc tn19_screeningvisitassessment.Ethnicity
Autoantibody positive	tn19_treatmentstartdate.TreatmentDesc tn19_researchlabs.Test_Name tn19_researchlabs.Visit tn19_researchlabs.result
Weight (kg)	tn19_treatmentstartdate.TreatmentDesc tn19_screeningvisitassessment.anthopometricsWeightKg
BMI (kg/m^2)	tn19_treatmentstartdate.TreatmentDesc tn19_screeningvisitassessment.AnthopometricsWeightKg tn19_screeningvisitassessment.AnthopometricsHeightCM
HbA1c	tn19_treatmentstartdate.TreatmentDesc tn19_researchlabs.test_name tn19_researchlabs.result

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

Characteristics	**ATG/GCSF (n=29)**	**DSIC ATG/GCSF (n=29)**	**Diff. (n=0)**	**ATG only (n=29)**	**DSIC ATG only (n=29)**	**Diff. (n=0)**	**Placebo (n=31)**	**DSIC Placebo (n=31)**	**Diff. (n=0)**
Age (years)									
Mean ± SD	17.2±5.0	17.1±5.0	0.1±0.0	18.1 ± 6.9	18.1±6.9	0±0	16.9 ± 4.6	16.8±4.6	0.1±0
Median	16.4	16.4	0	15.5	15.5	0	15.0	14.9	.1
Range	12.0-32.8	12.0-32.7	0-0.1	12.4-42.5	12.4-42.5	0-0	12.2-29.3	12.1-29.2	0.1-0.1
Male Sex	16 (55.2)	16 (55.2)	0(0.0)	17 (58.6)	17 (58.6)	0 (0.0)	17 (54.8)	17 (54.8)	0 (0.0)
Race									
White	28 (96.6)	28 (96.6)	0 (0.0)	29 (100)	29 (100)	0 (0.0)	29 (93.5)	29 (93.5)	0 (0.0)
Black	1 (3.4)	1 (3.4)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (6.5)	2 (6.5)	0 (0.0)
Ethnicity									
Not Hispanic and not Latino	28 (96.6)	28 (96.6)	0 (0.0)	27 (93.1)	27 (93.1)	0 (0.0)	30 (96.8)	30 (96.8)	0 (0.0)
Autoantibody positive									
GAD-65H	23 (79.3)	23 (79.3)	0 (0.0)	23 (79.3)	23 (79.3)	0 (0.0)	23 (74.2)	23 (74.2)	0 (0.0)
IA-2H	25 (86.2)	25 (86.2)	0 (0.0)	23 (79.3)	23 (79.3)	0 (0.0)	25 (80.6)	24 (77.4)	1 (3.2)
ICA	26 (92.9)	26 (92.9)	0 (0.0)	25 (86.2)	25 (86.2)	0 (0.0)	22 (71.0)	22 (71.0)	0 (0.0)
ZnT8	21 (72.4)	21 (72.4)	0 (0.0)	21 (72.4)	21 (72.4)	0 (0.0)	22 (71.0)	22 (71.0)	0 (0.0)
Weight (kg)									
Median	62.3	62.3	0	66.4	66.4	0	62.0	62.0	0
Range	39.8-89.1	37.3-89.1	2.5-0	39.6-92.4	39.6-92.4	0-0	33.8-118	33.8-118	0-0
BMI (kg/m^2)									
Median	21.4	21.1	0.3	22.6	22.6	0	21.8	21.8	0
Range	16.6-27.7	16.6-27.7	0-0	15.2-32.8	15.2-32.8	0-0	14.3-34.3	14.3-34.3	0-0
HbA1c									
Median %	7.3	7.3	0	7.4	7.4	0	7.2	7.2	0
Median mmol/mol	56	56.3	0.3	57	57.3	0.3	55	55.2	0.2
Range %	5.3-12.3	5.3-12.3	0-0	4.7-9.0	4.7-9.0	0-0	5.5-11.2	5.5-11.2	0-0
Range mmol/mol	34-111	34-111	0-0	28-75	28-75	-0	35-99	37-100	2-1

Attachment A: SAS Code

```
libname dsic "X:\NIDDK\niddk-dr_studies6\TrialNet_19\private_created_data\DSIC Data\sasv9";

*DSIC for TN19 complete data;

*create temp datasets;
data screen; set dsic.tn19_screeningvisitassessment;
run;

data treat; set dsic.tn19_treatmentstartdate;
run;

data labs; set dsic.tn19_researchlabs;
run;

*Age for publication;
proc freq data=screen;
tables DateOfBirthMonth;
run;

data screen1; set screen;
if DateOfBirthMonth = "Jan" then birth_month = 1;
if DateOfBirthMonth = "Feb"      then birth_month = 2;
if DateOfBirthMonth = "Mar"      then birth_month = 3;
if DateOfBirthMonth = "Apr"      then birth_month = 4;
if DateOfBirthMonth = "May"      then birth_month = 5;
if DateOfBirthMonth = "Jun"      then birth_month = 6;
if DateOfBirthMonth = "Jul"      then birth_month = 7;
if DateOfBirthMonth = "Aug"      then birth_month = 8;
if DateOfBirthMonth = "Sep"      then birth_month = 9;
if DateOfBirthMonth = "Oct"      then birth_month = 10;
if DateOfBirthMonth = "Nov"      then birth_month = 11;
if DateOfBirthMonth = "Dec"      then birth_month = 12;
run;

*creating a single month/year dob variable;
data screen2; set screen1;
dob = mdy(birth_month,1,DateOfBirthYear);
run;

*creating age variable for publication;
data screen3; set screen2;
age = (Visit_Dt - dob)/365.25;
run;

proc freq data=screen3;
```

```
tables age;
run;

*merging treatment assignement dataset with screen;
proc sort data=screen3;
by MaskID;
run;

proc sort data=treat;
by MaskID;
run;

data screen4;
merge
screen3
treat;
by maskid;
run;

*age for DSIC;
proc sort data=screen4;
by TreatmentDesc;
run;

proc means data=screen4 n mean std median min max;
var age;
by TreatmentDesc;
run;

*male sex;
proc freq data=screen4;
tables SexGenderPerson*TreatmentDesc/norow nopercent;
run;

*Race;
data screen5; set screen4;
Race = 0;
if Race_white = 1 then race = 1;
if race_blackorafricanamerican = 1 then race = 2;
run;

proc freq data=screen5;
tables race*TreatmentDesc/norow nopercent;
run;

*Ethnicity;
proc freq data=screen5;
tables Ethnicity*TreatmentDesc/norow nopercent;
```

```
run;

*autoantibodies;
proc freq data=labs;
tables Test_Name Event_Title Spec_Name;
run;

proc freq data=labs;
tables Visit;
run;

data labs1; set labs;
if visit = "Screening";
run;

data labs2; set labs1;
if test_name = "GAD65H" or test_name = "IA-2H" or test_name = "ICA" or test_name = "ZNT8";
run;

*getting positive and negative results based on lab result values;
data labs3; set labs2;
if test_name = "GAD65H" AND result > 20 then GAD = 1; else GAD = 0;
if test_name = "IA-2H" AND result > 5 then IA2H = 1; else IA2h = 0;
if test_name = "ICA" AND result >= 10 then ICA = 1; else ICA = 0;
if test_name = "ZNT8" AND result > 0.02 then ZNT8 = 1; else ZNT8 = 0;
run;

*merging labs and treatment datasets;
proc sort data=treat;
by MaskID;
run;

proc sort data=labs3;
by MaskID;
run;

data labs4;
merge
treat
labs3;
by maskid;
run;

*AB results;
proc freq data=labs4;
tables gad*TreatmentDesc/norow nopercent;
run;
```

```
proc freq data=labs4;
tables ia2h*TreatmentDesc/norow nopercent;
run;

proc freq data=labs4;
tables ica*TreatmentDesc/norow nopercent;
run;

proc freq data=labs4;
tables znt8*TreatmentDesc/norow nopercent;
run;

*weight;
proc means data=screen5 n median min max;
var AnthopometricsWeightKg;
by TreatmentDesc;
run;

*BMI;
data screen6; set screen5;
heightm = anthopometricsHeightCM/100;
bmi = anthopometricsWeightKg/(heightm * heightm);
run;

proc means data=screen6 n median min max;
var bmi;
by TreatmentDesc;
run;

*HbA1c;
data hba1c; set labs;
if Test_Name = "HbA1c" and visit = "Screening";
run;

*mergins Hba1c with treatment dataset;
proc sort data=hba1c;
by MaskID;
run;

data hba1c1;
merge
hba1c
treat;
by maskid;
run;

*converting char variable to numeric;
data hba1c2; set hba1c1;
```

```
result1 = input(result, 5.);
run;

proc sort data=hba1c2;
by TreatmentDesc;
run;

*HbA1c % results;
proc means data=hba1c2 n median min max;
var result1;
by TreatmentDesc;
run;

*creating a HbA1c mmol/mol variable;
data hba1c3; set hba1c2;
hba1c = 10.929 * (result1 - 2.15);
run;

proc means data=hba1c3 n median min max;
var hba1c;
by TreatmentDesc;
run;
```