

Dataset Integrity Check for The Environmental Determinants of Diabetes in the Young (TEDDY) M158 Mattila

Prepared by NIDDK-CR
September 17, 2021

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 – Mean childhood plasma ascorbic acid in islet autoimmunity and type 1 diabetes cases and controls	4
Table B1: Comparison of values computed in integrity check to reference article Table 1 values (Islet autoimmunity)	5
Table B2: Comparison of values computed in integrity check to reference article Table 1 values (Type 1 diabetes)	6
Attachment A: SAS Code	7

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 type 1 diabetes (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.

The M158 study sought to determine whether single nucleotide polymorphisms (SNPs) are associated with the risk of islet autoimmunity or type 1 diabetes.

3 Archived Datasets

All 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_158_ierlund_niddk_31may2012_2.sas7bdat”, and “m_158_ierlund_niddk_31may2012_3.sas7bdat” datasets.

4 Statistical Methods

Analyses were performed to replicate results for the data published by Mattila et al. [1] for Plasma Ascorbic Acid and the Risk of Islet Autoimmunity and Type 1 Diabetes: the TEDDY Study. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Mean childhood plasma ascorbic acid in islet autoimmunity and type 1 diabetes cases and controls, Table A lists the variables that were used in the replication, and Tables B1-B2 compare the results calculated from the archived data files to the results published in Table 1. The results of the replication are within expected variation to the published results.

6 Conclusions

The NIDDK Central Repository is confident that the TEDDY M158 data files to be distributed are a true copy of the study data.

7 References

[1] Mattila M, Erlund I, Lee HS, Niinistö S, Uusitalo U, Aronsson CA, Hummel S, Parikh H, Rich SS, Hagopian W, Toppari J, Lernmark Å, Ziegler AG, Rewers M, Krischer JP, Norris JM, Virtanen SM. Plasma Ascorbic Acid and the Risk of Islet Autoimmunity and Type 1 Diabetes: the TEDDY Study. *Diabetologia*, 63(2), 278-286, 2020. doi: <https://doi.org/10.1007/s00125-019-05028-z>

Table A: Variables used to replicate Table 1 – Mean childhood plasma ascorbic acid in islet autoimmunity and type 1 diabetes cases and controls

Table Variable	dataset.variable
Clinical centre	m_158_ierlund_niddk_31may2012_2.cc m_158_ierlund_niddk_31may2012_3.cc
Sex	m_158_ierlund_niddk_31may2012_2.sex m_158_ierlund_niddk_31may2012_3.sex
FDR/GP status	m_158_ierlund_niddk_31may2012_2.fdr m_158_ierlund_niddk_31may2012_3.fdr

Table B1: Comparison of values computed in integrity check to reference article Table 1 values (Islet autoimmunity)

Matching variable	Islet autoimmunity								
	No. (%) of cases			Plasma ascorbic acid concentration (mg/l)					
				Cases			Controls		
	Publication	DSIC	Diff.	Publication	DSIC	Diff.	Publication	DSIC	Diff.
Clinical centre									
Colorado	51 (14.6)	51 (14.6)	0 (0)	11.7 ± 3.0	11.7 ± 3.0	0 (0)	12.2 ± 3.3	12.2 ± 3.3	0 (0)
Georgia	24 (6.9)	24 (6.9)	0 (0)	12.5 ± 3.5	12.5 ± 3.5	0 (0)	12.3 ± 3.4	12.3 ± 3.4	0 (0)
Washington State	34 (9.7)	34 (9.7)	0 (0)	11.4 ± 4.3	11.3 ± 4.3	0.1 (0)	11.8 ± 4.4	11.8 ± 4.3	0 (0.1)
Finland	105 (30.0)	105 (30.0)	0 (0)	10.5 ± 2.8	10.5 ± 2.8	0 (0)	10.7 ± 3.0	10.7 ± 3.0	0 (0)
Germany	26 (7.4)	26 (7.4)	0 (0)	9.2 ± 2.5	9.2 ± 2.5	0 (0)	9.7 ± 3.4	9.7 ± 3.4	0 (0)
Sweden	110 (31.4)	110 (31.4)	0 (0)	8.7 ± 3.0	8.7 ± 3.0	0 (0)	9.5 ± 3.6	9.5 ± 3.6	0 (0)
Sex									
Female	157 (44.9)	157 (44.9)	0 (0)	10.0 ± 3.4	10.0 ± 3.4	0 (0)	10.7 ± 3.5	10.7 ± 3.5	0 (0)
Male	193 (55.1)	193 (55.1)	0 (0)	10.4 ± 3.2	10.4 ± 3.2	0 (0)	10.8 ± 3.7	10.8 ± 3.7	0 (0)
FDR/GP status									
FDR	76 (21.7)	76 (21.7)	0 (0)	10.8 ± 3.1	10.8 ± 3.1	0 (0)	11.1 ± 3.7	11.1 ± 3.8	0 (0.1)
GP	274 (78.3)	274 (78.3)	0 (0)	10.0 ± 3.4	10.0 ± 3.4	0 (0)	10.7 ± 3.5	10.7 ± 3.5	0 (0)

Table B2: Comparison of values computed in integrity check to reference article Table 1 values (Type 1 diabetes)

Matching variable	Type 1 diabetes								
	No. (%) of cases			Plasma ascorbic acid concentration (mg/l)					
	Publication	DSIC	Diff.	Cases			Controls		
Publication				DSIC	Diff.	Publication	DSIC	Diff.	
Clinical centre									
Colorado	15 (14.7)	15 (14.7)	0 (0)	12.0 ± 2.5	12.0 ± 2.5	0 (0)	12.2 ± 3.1	12.2 ± 3.1	0 (0)
Georgia	6 (5.9)	6 (5.9)	0 (0)	13.1 ± 4.2	13.1 ± 4.2	0 (0)	13.6 ± 4.1	13.6 ± 4.1	0 (0)
Washington State	7 (6.9)	7 (6.9)	0 (0)	9.8 ± 3.5	9.8 ± 3.5	0 (0)	11.6 ± 3.2	11.6 ± 3.2	0 (0)
Finland	35 (34.3)	35 (34.3)	0 (0)	9.8 ± 2.7	9.8 ± 2.7	0 (0)	10.7 ± 3.6	10.7 ± 3.6	0 (0)
Germany	15 (14.7)	15 (14.7)	0 (0)	10.2 ± 3.2	10.1 ± 3.2	0.1 (0)	10.3 ± 3.8	10.3 ± 3.7	0 (0.1)
Sweden	24 (23.5)	24 (23.5)	0 (0)	7.9 ± 0.9	7.9 ± 2.9	0 (2.0)	8.8 ± 3.0	8.8 ± 3.0	0 (0)
Sex									
Female	47 (46.1)	47 (46.1)	0 (0)	10.2 ± 3.1	10.2 ± 3.1	0 (0)	10.2 ± 3.1	10.2 ± 3.1	0 (0)
Male	55 (53.9)	55 (53.9)	0 (0)	9.7 ± 3.3	9.7 ± 3.3	0 (0)	11.2 ± 4.0	11.2 ± 4.0	0 (0)
FDR/GP status									
FDR	36 (35.3)	36 (35.3)	0 (0)	10.8 ± 2.8	10.8 ± 2.8	0 (0)	11.0 ± 3.8	11.0 ± 3.8	0 (0)
GP	66 (64.7)	66 (64.7)	0 (0)	9.5 ± 3.4	9.5 ± 3.4	0 (0)	10.6 ± 3.6	10.6 ± 3.6	0 (0)

Attachment A: SAS Code

```
libname dsic "X:\NIDDK\niddk-
dr_studies6\TEDDY\private_orig_data\M_158_I Erlund_NIDDK_Submission";

/*****
/* Dataset Integrity Check (DSIC) for the */
/* "Plasma ascorbic acid and the risk of islet autoimmunity and type 1 diabetes: the TEDDY study" */
/* Mattila et al. */
*****/

*Temporary datasets for analysis;
data m158_1; set dsic.m_158_ierlund_niddk_31may2012_1;
run;

data m158_2; set dsic.m_158_ierlund_niddk_31may2012_2;
run;

data m158_3; set dsic.m_158_ierlund_niddk_31may2012_3;
run;

data m158_4; set dsic.m_158_ierlund_niddk_31may2012_4;
run;

*Checking the contents of the datasets;
proc contents data=m158_1;
run;

proc contents data=m158_2;
run;

proc contents data=m158_3;
run;

proc contents data=m158_4;
run;

/*****
/* Islet Autoimmunity */
*****/

*Number of Cases (%): Clinical Centre;
proc freq data=m158_1;
tables cc*outcome/norow nopercnt;
run;

*Plasma ascorbic acid concentration(mg/l): Cases and Controls by Clinical Centre;
proc sort data=m158_2;
by outcome;
run;

proc means data=m158_2 mean std;
var avVC;
class cc;
by outcome;
```



```

run;

*Number of Cases (%): Sex;
proc freq data=m158_2;
tables sex*outcome/norow nopercnt;
run;

*Plasma ascorbic acid concentration(mg/l): Cases and Controls by Sex;
proc means data=m158_2 mean std;
var avVC;
class sex;
by outcome;
run;

*Number of Cases (%): FDR/GP Status;
proc freq data=m158_2;
tables fdr*outcome/norow nopercnt;
run;

*Plasma ascorbic acid concentration (mg/l): Cases and Controls by FDR/GP Status;
proc means data=m158_2 mean std;
var avVC;
class fdr;
by outcome;
run;

/*****/
/* Type 1 Diabetes */
/*****/

*Number of cases (%): Clinical Centre;
proc freq data=m158_3;
tables cc*outcome/norow nopercnt;
run;

*Plasma ascorbic acid concentration (mg/l): Cases and controls by Clinical Centre;
proc sort data=m158_3;
by outcome;
run;

proc means data=m158_3 mean std;
var avVC;
class cc;
by outcome;
run;

*Number of cases (%): Sex;
proc freq data=m158_3;
tables sex*outcome/norow nopercnt;
run;

*Plasma ascorbic acid concentration (mg/l): Cases and COntrols by Sex;
proc means data=m158_3 mean std;

```

```
var avVC;  
class sex;  
by outcome;  
run;
```

```
*Number of Cases (%): FDR/GP Status;  
proc freq data=m158_3;  
tables fdr*outcome/norow nopercnt;  
run;
```

```
*Plasma ascorbic acid concentration (mg/l): Cases and Controls by FDR/GP Status;  
proc means data=m158_3 mean std;  
var avVC;  
class fdr;  
by outcome;  
run;
```