Dataset Integrity Check for The Environmental Determinants of Diabetes in the Young (TEDDY) M179 Vehik

Prepared by NIDDK-CR
August 17, 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 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. Data are 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 M179 study sought to assess the risk of first-appearing beta-cell autoantibody spreading to the second appearing autoantibody and subsequent progression to clinical disease in the TEDDY cohort.

3 Archived Datasets

All 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_179_kvehik_niddk_31dec2019_1.sas7bdat", "m_179_kvehik_niddk_31dec2019_2.sas7bdat", and "m_179_kvehik_niddk_31dec2019_3.sas7bdat" datasets.

4 Statistical Methods

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

5 Results

For Supplemental Table S2 in the publication [1], Supplemental Table S2 – Characteristics of TEDDY children who developed a distinct second-appearing autoantibody by first-appearing IAA or GADA (n=608), Table A lists the variables that were used in the replication, and Tables B1 and B2 compare the results calculated from the archived data files to the results published in Supplemental Table S2. The results of the replication are within expected variation to the published results. Minor allele frequencies were not calculated in this replication.

6 Conclusions

The NIDDK Central Repository is confident that the TEDDY M179 data files to be distributed are a true copy of the study data.

7 References

[1] Vehik K, Bonifacio E, Lernmark Å, Yu L, Williams A, Schatz D, Rewers M, She JX, Toppari J, Hagopian W, Akolkar B, Ziegler AG, Krischer JP. Hierarchical Order of Distinct Autoantibody Spreading and Progression to Type 1 Diabetes in the TEDDY Study. Diabetes Care, 43(9), 2066-2073, July 2020. doi: https://doi.org/10.2337/dc19-2547

Table A: Variables used to replicate Supplemental Table S2 – Characteristics of TEDDY children who developed a distinct second-appearing autoantibody by first-appearing IAA or GADA (n=608)

Table Variable	**dataset.variable**
n (%)	m179.m_179_kvehik_niddk_31dec2019_2.group
Sex, n (% female)	m179.m_179_kvehik_niddk_31dec2019_2.sex
Family History with T1D, n (% yes)	m179.m_179_kvehik_niddk_31dec2019_2.family_screen
HLA-DR-DQ genotype, n(%)	m179.m_179_kvehik_niddk_31dec2019_2.hla_5grps
HLA A*24:X, n(%)	m179.m_179_kvehik_niddk_31dec2019_2.hlaA24
HLA B*18:X, n(%)	m179.m_179_kvehik_niddk_31dec2019_2.hlaB18
HLA B*39:X, n(%)	m179.m_179_kvehik_niddk_31dec2019_2.hlaB39
Maternal islet autoimmunity, n (% yes)	m179.m_179_kvehik_niddk_31dec2019_2.maternal_ab_exp
Probiotics started by 28 days, n (% yes)	m179.m_179_kvehik_niddk_31dec2019_2.probiotics_by_28days
Weight at 12 months, median z-score (IQR)	m179.m_179_kvehik_niddk_31dec2019_2.waz
Child conditions before first clinical visit at age 3 months, n (%)	m179.m_179_kvehik_niddk_31dec2019_2.child_upper_resp_3mo m179.m_179_kvehik_niddk_31dec2019_2.child_lower_resp_3mo m179.m_179_kvehik_niddk_31dec2019_2.child_diarrhea_episode m179.m_179_kvehik_niddk_31dec2019_2.child_rash
Median (IQR) Age at Second Appearing Autoantibody (months)	m179.m_179_kvehik_niddk_31dec2019_2.persist_conf_second_ab_mos
Median (IQR) duration of time from First to Second-Appearing Autoantibody (months)	m179.m_179_kvehik_niddk_31dec2019_2.timefrom_persist_mos
Progression to type 1 diabetes from time of second appearing autoantibody, n (%)	m179.m_179_kvehik_niddk_31dec2019_2.t1d
Median (IQR) time from Second Appearing Autoantibody to type 1 diabetes (months)	m179.m_179_kvehik_niddk_31dec2019_2.timefrom_secondAb_mos m179.m_179_kvehik_niddk_31dec2019_2.timefrompersistt1d m179.m_179_kvehik_niddk_31dec2019_2.timefrom_persist_mos
Developed a third or fourth autoantibody, n (% yes)	m179.m_179_kvehik_niddk_31dec2019_2.third_ab
Median (IQR) age of follow-up (months) from initial seroconversion	m179.m_179_kvehik_niddk_31dec2019_2.persist_conf_second_ab_mos m179.m_179_kvehik_niddk_31dec2019_2.persist_atrisk_age m179.m_179_kvehik_niddk_31dec2019_2.t1d_atrisk_agemos m179.m_179_kvehik_niddk_31dec2019_2.timefrom_persist_mos m179.m_179_kvehik_niddk_31dec2019_2.timefrom_secondAb_mos m179.m_179_kvehik_niddk_31dec2019_2.timefrompersistt1d

Table B1: Comparison of values computed in integrity check to reference article Supplemental Table S2 (IAA-First appearing)

Publication: IAA-First appearing children at risk for second-appearing (n=282)							DSIC: IAA-First appearing children at risk for second-appearing (n=282)					Diff. (n=0)
Second appearing autoantibody	**Pub: GADA**	**DSIC: GADA**	**Diff.**	**Pub: IA-2A**	**DSIC: IA-2A**	**Diff.**	**Pub: ZnT8A**	**DSIC: ZnT8A**	**Diff.**	**Pub: None**	**DSIC: None**	**Diff.**
N (%)	85 (30.1)	85 (30.1)	0 (0)	31 (11.0)	31 (11.0)	0 (0)	23 (8.2)	23 (8.2)	0 (0)	116 (41.1)	116 (41.1)	0 (0)
Sex, n (% female)	35 (41.2)	35 (41.2)	0 (0)	14 (45.2)	14 (45.2)	0 (0)	7 (30.4)	7 (30.4)	0 (0)	59 (50.9)	59 (50.9)	0 (0)
Family History with T1D, n (% yes)												
Sibling with T1D	9 (10.6)	9 (10.6)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	0	0	0	4 (3.5)	4 (3.5)	0 (0)
Father with T1D	10 (11.8)	10 (11.8)	0 (0)	4 (12.9)	4 (12.9)	0 (0)	5 (21.7)	5 (21.7)	0 (0)	5 (4.3)	5 (4.3)	0 (0)
Mother with T1D	2 (2.4)	2 (2.4)	0 (0)	2 (6.5)	2 (6.5)	0 (0)	0	0	0	3 (3.0)	3 (2.6)	0 (0.4)
HLA-DR-DQ genotype, n (%)												
DR3/4	53 (62.4)	53 (62.4)	0 (0)	12 (38.7)	12 (38.7)	0 (0)	14 (60.9)	14 (60.9)	0 (0)	46 (39.7)	46 (39.7)	0 (0)
DR4/4	18 (21.2)	18 (21.2)	0 (0)	6 (19.4)	6 (19.4)	0 (0)	3 (13.0)	3 (13.0)	0 (0)	16 (13.8)	16 (13.8)	0 (0)
DR4/8	8 (9.4)	8 (9.4)	0 (0)	12 (38.7)	12 (38.7)	0 (0)	5 (21.7)	5 (21.7)	0 (0)	29 (25.0)	29 (25.0)	0 (0)
DR3/3	2 (2.4)	2 (2.4)	0 (0)	0	0	0	1 (4.4)	1 (4.4)	0 (0)	20 (17.2)	20 (17.2)	0 (0)
First Degree Relative (FDR)	4 (4.7)	4 (4.7)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	0	0	0	5 (4.3)	5 (4.3)	0 (0)
HLA A*24:X, n (%)	11 (12.9)	11 (12.9)	0 (0)	9 (29.0)	9 (29.0)	0 (0)	11 (47.8)	4 (17.4)	7 (30.4)	19 (16.4)	19 (16.4)	0 (0)
HLA B*18:X, n (%)	13 (15.3)	13 (15.3)	0 (0)	5 (16.1)	5 (16.1)	0 (0)	1 (4.4)	1 (4.4)	0 (0)	20 (17.2)	20 (17.2)	0 (0)
HLA B*39:X, n (%)	5 (5.9)	5 (5.9)	0 (0)	5 (16.1)	5 (16.1)	0 (0)	2 (8.7)	2 (8.7)	0 (0)	7 (6.0)	7 (6.0)	0 (0)
Maternal islet autoimmunity, n (% yes)	3 (3.5)	3 (3.5)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	0	0	0	2 (1.7)	2 (1.7)	0 (0)
Probiotics started by 28 days, n (% yes)	4 (4.7)	4 (4.7)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	3 (13.0)	3 (13.0)	0 (0)	5 (4.3)	5 (4.3)	0 (0)
Weight at 12 months, median z-score (IQR)	-0.13 (-1.00-0.67)	-0.13 (-1.00-0.67)	0 (0-0)	-0.35 (-0.91-0.45)	-0.35 (-0.91-0.45)	0 (0-0)	-0.10 (-0.77-0.57)	-0.10 (-0.77-0.57)	0 (0-0)	-0.06 (-0.75-0.71)	-0.06 (-0.75-0.71)	0 (0-0)
Child conditions before 1st clinical visit at age 3 months, n (%)												
Upper respiratory infection	20 (23.5)	20 (23.5)	0 (0)	6 (19.4)	6 (19.4)	0 (0)	6 (26.1)	6 (26.1)	0 (0)	27 (23.3)	27 (23.3)	0 (0)
Lower respiratory infection	14 (16.5)	14 (16.5)	0 (0)	3 (9.7)	3 (9.7)	0 (0)	0	0	0	11 (9.5)	11 (9.5)	0 (0)
Diarrhea	7 (8.2)	7 (8.2)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	3 (13.0)	3 (13.0)	0 (0)	14 (12.1)	14 (12.1)	0 (0)
Rash	16 (18.9)	16 (18.8)	0 (0.1)	3 (9.7)	3 (9.7)	0 (0)	7 (22.6)	7 (30.4)	0 (7.8)	24 (20.7)	24 (20.7)	0 (0)
Median (IQR) Age at second Appearing Autoantibody (months)	29.9 (18.5-45.0)	29.9 (18.5-45.0)	0 (0-0)	19.4 (17.8-29.2)	19.4 (17.8-29.2)	0 (0-0)	37.2 (21.1-59.8)	37.2 (21.1-59.8)	0 (0-0)	NA	NA	NA

Publication: IAA-First appearing children at risk for second-appearing (n=282)							DSIC: IAA-First appearing children at risk for second-appearing (n=282)					Diff. (n=0)
Second appearing autoantibody	**Pub: GADA**	**DSIC: GADA**	**Diff.**	**Pub: IA-2A**	**DSIC: IA-2A**	**Diff.**	**Pub: ZnT8A**	**DSIC: ZnT8A**	**Diff.**	**Pub: None**	**DSIC: None**	**Diff.**
Median (IQR) duration of time from First to Second-Appearing Autoantibody (months)	3.7 (3.0-9.7)	3.7 (3.0- 9.7)	0 (0-0)	5.9 (3.2-9.3)	5.9 (3.2-9.3)	0 (0-0)	9.6 (3.3-24.1)	9.6 (3.3-24.1)	0 (0-0)	NA	NA	NA
Progression to type 1 diabetes from time of second appearing autoantibody, n (%)	49 (57.7)	49 (57.7)	0 (0)	28 (90.3)	28 (90.3)	0 (0)	11 (47.8)	11 (47.8)	0 (0)	18 (15.5)	18 (15.5)	0 (0)
Median (IQR) time from Second Appearing Autoantibody to type 1 diabetes (months)	51.0 (20.1-76.8)	51.0 (20.1-76.8)	0 (0-0)	27.2 (9.2-43.2)	28.4 (12.5-55.2)	1.2 (3.3-12)	75.9 (38.2-92.1)	75.8 (42.5-88.7)	0.1 (4.3-3.4)	6.7 (1.7-23.9)	5.4 (2.8-25.0)	1.3 (1.1-1.1)
Developed a third or fourth autoantibody, n (% yes)	68 (80.0)	68 (80.0)	0 (0)	17 (54.8)	17 (54.8)	0 (0)	17 (73.9)	17 (73.9)	0 (0)	NA	NA	NA
Median (IQR) age of follow-up (months) from initial seroconversion	82.3 (42.8-114.5)	82.3 (42.8-114.5)	0 (0-0)	34.9 (15.4-63.1)	34.9 (15.4-63.1)	0 (0-0)	109.6 (67.1-138.8)	109.6 (67.1-138.8)	0 (0-0)	71.3 (38.3-111.1)	71.3 (38.3-111.1)	0 (0-0)

Table B2: Comparison of values computed in integrity check to reference article Supplemental Table S2 (GADA-First appearing)

Publication: GADA-First appearing children at risk for second-appearing (n=326)							DSIC: GADA-First appearing children at risk for second-appearing (n=326)					Diff. (n=0)
Second appearing autoantibody	**Pub: IAA**	**DSIC: IAA**	**Diff.**	**Pub: IA-2A**	**DSIC: IA-2A**	**Diff.**	**Pub: ZnT8A**	**DSIC: ZnT8A**	**Diff.**	**Pub: None**	**DSIC: None**	**Diff.**
N (%)	76 (23.1)	76 (23.1)	0 (0)	31 (9.5)	31 (9.5)	0 (0)	43 (13.2)	43 (13.2)	0 (0)	156 (47.9)	156 (47.9)	0 (0)
Sex, n (% female)	34 (44.7)	34 (44.7)	0 (0)	17 (54.2)	17 (54.8)	0 (0.6)	17 (39.5)	17 (39.5)	0 (0)	78 (50.0)	78 (50.0)	0 (0)
Family History with T1D, n (% yes)												
Sibling with T1D	1 (1.3)	1 (1.3)	0 (0)	0	0	0	3 (7.1)	3 (7.0)	0 (0.1)	2 (1.3)	2 (1.3)	0 (0)
Father with T1D	9 (11.8)	9 (11.8)	0 (0)	5 (15.6)	4 (12.9)	1 (2.7)	4 (9.5)	4 (9.3)	0 (0.2)	11 (7.1)	11 (7.1)	0 (0)
Mother with T1D	3 (4.0)	3 (4.0)	0 (0)	3 (9.4)	3 (9.7)	0 (0.3)	3 (7.1)	3 (7.0)	0 (0.1)	6 (3.9)	6 (3.9)	0 (0)
HLA-DR-DQ genotype, n(%)												
DR3/4	47 (61.8)	47 (61.8)	0 (0)	16 (51.6)	16 (51.6)	0 (0)	19 (44.2)	19 (44.2)	0 (0)	69 (44.2)	69 (44.2)	0 (0)
DR4/4	11 (14.5)	11 (14.5)	0 (0)	4 (12.9)	4 (12.9)	0 (0)	6 (14.0)	6 (14.0)	0 (0)	25 (16.0)	25 (16.0)	0 (0)
DR4/8	9 (11.8)	9 (11.8)	0 (0)	4 (12.9)	4 (12.9)	0 (0)	5 (11.6)	5 (11.6)	0 (0)	21 (13.5)	21 (13.5)	0 (0)
DR3/3	8 (10.5)	8 (10.5)	0 (0)	7 (22.6)	7 (22.6)	0 (0)	13 (30.2)	13 (30.2)	0 (0)	39 (25.0)	39 (25.0)	0 (0)
First Degree Relative (FDR)	1 (1.3)	1 (1.3)	0 (0)	0	0	0	0	0	0	2 (1.3)	2 (1.3)	0 (0)
HLA A*24:X, n (%)	11 (14.5)	11 (14.5)	0 (0)	5 (16.1)	5 (16.1)	0 (0)	4 (9.3)	4 (9.3)	0 (0)	29 (18.6)	29 (18.6)	0 (0)
HLA B*18:X, n (%)	7 (9.2)	7 (9.2)	0 (0)	4 (12.9)	4 (12.9)	0 (0)	8 (18.6)	8 (18.6)	0 (0)	13 (8.3)	13 (8.3)	0 (0)
HLA B*39:X, n (%)	3 (4.0)	3 (4.0)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	4 (9.3)	4 (9.3)	0 (0)	5 (3.2)	5 (3.2)	0 (0)
Maternal islet autoimmunity, n(% yes)	1 (1.3)	1 (1.3)	0 (0)	1 (3.2)	1 (3.2)	0 (0)	0	0	0	5 (2.1)	5 (3.2)	0 (1.1)
Probiotics started by 28 days, n (% yes)	3 (4.0)	3 (4.0)	0 (0)	3 (9.7)	3 (9.7)	0 (0)	3 (7.0)	3 (7.0)	0 (0)	8 (5.1)	8 (5.1)	0 (0)
Weight at 12 months, median z-score (IQR)	0.23 (-0.72-0.76)	0.23 (-0.72-0.76)	0 (0-0)	0.16 (-0.32-0.81)	0.16 (-0.32-0.81)	0 (0-0)	0.37 (-0.55-0.86)	0.37 (-0.55-0.86)	0 (0-0)	-0.005 (-0.63-0.66)	-0.005 (-0.63-0.66)	0 (0-0)
Child conditions before 1st clinical visit at age 3 months, n (%)												
Upper respiratory infection	12 (15.8)	12 (15.8)	0 (0)	8 (25.8)	8 (25.8)	0 (0)	8 (19.1)	8 (19.1)	0 (0)	34 (21.8)	34 (21.8)	0 (0)
Lower respiratory infection	11 (14.5)	11 (14.5)	0 (0)	5 (16.1)	5 (16.1)	0 (0)	7 (16.2)	7 (16.3)	0 (0.1)	10 (6.4)	10 (6.4)	0 (0)
Diarrhea	4 (5.2)	4 (5.3)	0 (0.1)	2 (6.5)	2 (6.5)	0 (0)	2 (4.7)	2 (4.7)	0 (0)	13 (8.3)	13 (8.3)	0 (0)
Rash	10 (13.2)	10 (13.2)	0 (0)	7 (22.6)	7 (22.6)	0 (0)	8 (18.6)	8 (18.6)	0 (0)	35 (22.4)	35 (22.4)	0 (0)
Median (IQR) Age at second Appearing Autoantibody (months)	57.8 (29.4-102.8)	57.8 (29.4-102.8)	0 (0-0)	65.6 (35.9-105.9)	65.6 (35.9-105.9)	0 (0-0)	60.9 (48.7-85.5)	60.9 (48.7-85.5)	0 (0-0)	NA	NA	NA

Publication: GADA-First appearing children at risk for second-appearing (n=326)							DSIC: GADA-First appearing children at risk for second-appearing (n=326)					Diff. (n=0)
Second appearing autoantibody	**Pub: IAA**	**DSIC: IAA**	**Diff.**	**Pub: IA-2A**	**DSIC: IA-2A**	**Diff.**	**Pub: ZnT8A**	**DSIC: ZnT8A**	**Diff.**	**Pub: None**	**DSIC: None**	**Diff.**
Median (IQR) duration of time from First to Second-Appearing Autoantibody (months)	5.9 (3.0-25.9)	5.9 (3.0-25.9)	0 (0-0)	12.4 (6.2-23.7)	12.4 (6.2-23.7)	0 (0-0)	15.2 (8.7-23.7)	15.2 (8.7-23.7)	0 (0-0)	NA	NA	NA
Progression to type 1 diabetes from time of second appearing autoantibody, n(%)	28 (36.8)	28 (36.8)	0 (0)	16 (51.6)	16 (51.6)	0 (0)	17 (39.5)	17 (39.5)	0 (0)	10 (6.4)	10 (6.4)	0 (0)
Median (IQR) time from Second Appearing Autoantibody to type 1 diabetes (months)	44.9 (18.9-68.6)	44.9 (18.9-68.6)	0 (0-0)	31.9 (16.4-61.1)	28.4 (12.5-55.2)	3.5 (3.9-5.9)	75.8 (42.6-86.7)	75.8 (42.5-88.7)	0 (0.1-2)	27.9 (4.5-40.8)	11.4 (3.4-40.4)	16.5 (1.1-0.4)
Developed a third or fourth autoantibody, n (% yes)	50 (65.8)	50 (65.8)	0 (0)	19 (61.3)	13 (41.9)	6 (19.4)	30 (69.8)	22 (51.2)	8 (18.6)	NA	NA	NA
Median (IQR) age of follow-up (months) from initial seroconversion	65.4 (45.2-102.9)	65.4 (45.2-102.9)	0 (0-0)	65.0 (33.4-87.0)	65.0 (33.4-87.0)	0 (0-0)	91.9 (59.6-107.4)	91.9 (59.6-107.4)	0 (0-0)	63.3 (36.4-93.7)	63.3 (36.4-93.7)	0 (0-0)

Attachment A: SAS Code

```
libname m179 "X:\NIDDK\niddk-
dr_studies6\TEDDY\private_orig_data\M_179_KVehik_NIDDK_Submission";

/*********************************************/
/*       DSIC for TEDDY M179 - Vehik         */
/*********************************************/

*Calling datasets;

data one; set m179.m_179_kvehik_niddk_31dec2019_1;
run;

data two; set m179.m_179_kvehik_niddk_31dec2019_2;
run;

data three; set m179.m_179_kvehik_niddk_31dec2019_3;
run;

*reviewing the datasets;
proc contents data=one;
run;
proc contents data=two;
run;
proc contents data=three;
run;

/*replicating Table S2*/
/*IAA-First appearing children at risk for second-appearing*/

*IAA-First appearing children at risk for second-appearing;
proc freq data=two;
tables group; *(1=IAA only);
run;

*limiting the data to only those with IAA-first appearing;
*N(%) for second appearing ABs;
proc freq data=two;
tables group*secondab;
where group=1;
run;

*sex by second appearing AB;
proc freq data=two;
tables sex*secondab/norow nopercent;
```

```
where group=1;
run;

*Family history with T1D;
proc freq data=two;
tables family_screen*secondab/norow nopercent;
where group=1;
run;

*HLA-DR-DQ genotypes;
proc freq data=two;
tables hla_5grps*secondab/ norow nopercent;
where group=1;
run;

*HLA A*24;
proc freq data=two;
tables hlaA24*secondab /norow nopercent;
where group=1;
run;

*HLA B*18;
proc freq data=two;
tables hlaB18*secondab/norow nopercent;
where group=1;
run;

*HLA B*39;
proc freq data=two;
tables hlaB39*secondab/norow nopercent;
where group=1;
run;

*maternal islet AB;
proc freq data=two;
tables maternal_ab_exp*secondab/norow nopercent;
where group=1;
run;

*probiotics started by 28 days;
proc freq data=two;
tables probiotics_by_28days*secondab/norow nopercent;
where group=1;
run;

*weight at 12 months median z-score;
proc sort data=two;
by secondab;
```

```
run;

proc means data=two median q1 q3;
var waz;
by secondab;
where group=1;
run;

*child conditions before first clinical visit;
proc freq data=two;
tables (child_upper_resp_3mo child_lower_resp_3mo child_diarrhea_episode child_rash)*secondab
/norow nopercent missing;
where group=1;
run;

*SNPs minor allele frequency;
/*proc freq data=two;
tables (rs2476601_A rs1534422_G rs10517086_A rs2327832_G rs1004446_A rs2292239_A rs3184504_A
rs13266634_a_)*secondab/ norow nopercent missing;
where group=1;
run;
*/

*Median age at second appearing AB;
proc means data=two median q1 q3;
var persist_conf_second_ab_mos;
by secondab;
where group=1;
run;

*median duration of time from first to second appearing AB;
proc means data=two median q1 q3;
var timefrom_persist_mos;
by secondab;
where group = 1;
run;

*progresssion to type 1 diabetes from time of second appearing autoantibody;
proc freq data=two;
tables t1d*secondab/norow nopercent;
where group = 1;
run;

*median time from first and second appearing autoantibody to T1D;
proc means data=two median q1 q3;
var timefrom_secondAb_mos timefrompersistt1d timefrom_persist_mos;
by secondab;
where group = 1;
```

```
where t1d = 1;
run;

*developed a third or fourth AB;
data two_1; set two;
if third_ab = 1 OR third_ab = 3 OR third_ab = 4 then third_fourth_ab = 1; else third_fourth_ab = 0;
run;

proc freq data=two_1;
tables third_fourth_ab*secondab/norow nopercent;
where group = 1;
run;

*median age of follow-up from initial serconversion;
proc means data=two median q1 q3;
var persist_conf_second_ab_mos persist_atrisk_age T1D_atrisk_agemos timefrom_persist_mos
timefrom_secondAb_mos timefrompersistt1d;
by secondab;
where group = 1;
run;

/* GADA-First appeaing children at risk for second appearing*/
proc freq data=two;
tables group;
run;

*second appearing AB;
proc freq data=two;
tables group*secondab;
where group = 0;
run;

*Sex;
proc freq data=two;
tables sex*secondab/norow nopercent;
where group = 0;
run;

*family history with T1D;
proc freq data=two;
tables family_screen*secondab/norow nopercent;
where group = 0;
run;

*HLA-DR-DQ genotype;
proc freq data=two;
tables hla_5grps*secondab/norow nopercent;
where group = 0;
```

```
run;

*HLA A*24;
proc freq data=two;
tables hlaA24*secondab/norow nopercent;
where group=0;
run;

*HLA B*18;
proc freq data=two;
tables hlaB18*secondab/norow nopercent;
where group = 0;
run;

*HLA B*39;
proc freq data=two;
tables hlaB39*secondab/norow nopercent;
where group = 0;
run;

*Maternal islet immunity;
proc freq data=two;
tables maternal_ab_exp*secondab/norow nopercent;
where group = 0;
run;

*Probiotics started by 28 days;
proc freq data=two;
tables probiotics_by_28days*secondab/norow nopercent;
where group = 0;
run;

*weight at 12 months;
proc means data=two median q1 q3;
var waz;
by secondab;
where group = 0;
run;

*child conditions;
*upper resp inf;
proc freq data=two;
tables child_upper_resp_3mo*secondab/norow nopercent missing;
where group = 0;
run;

*lower resp inf;
proc freq data=two;
```

```
tables child_lower_resp_3mo*secondab/norow nopercent missing;
where group = 0;
run;

*Diarrhea;
proc freq data=two;
tables child_diarrhea_episode*secondab/norow nopercent missing;
where group = 0;
run;

*Rash;
proc freq data=two;
tables child_rash*secondab/norow nopercent missing;
where group = 0;
run;

*Median age at second appearing ab;
proc means data=two median q1 q3;
var persist_conf_second_ab_mos;
by secondab;
where group = 0;
run;

*median duration of time from first to second appearing ab;
proc means data=two median q1 q3;
var timefrom_persist_mos;
by secondab;
where group = 0;
run;

*progression to T1D from time of second appearing ab;
proc freq data=two;
tables t1d*secondab/norow nopercent missing;
where group = 0;
run;

*median time from second appearing ab to T1D;
proc means data=two median q1 q3;
var timefrom_secondAb_mos timefrompersistt1d timefrom_persist_mos;
by secondab;
where group = 0;
where t1d = 1;
run;

*developed a third or fourth ab;
data two_1; set two;
if third_ab = 1 OR third_ab = 3 OR third_ab = 4 then third_fourth_ab = 1; else third_fourth_ab = 0;
run;
```

```
proc freq data=two_1;
tables third_fourth_ab*secondab/norow nopercent;
where group = 0;
run;

*median age of follow up;
proc means data=two median q1 q3;
var persist_conf_second_ab_mos persist_atrisk_age T1D_atrisk_agemos timefrom_persist_mos
timefrom_secondAb_mos timefrompersistt1d;
by secondab;
where group = 0;
run;
```