Dataset Integrity Check for The Environmental Determinants of Diabetes in the Young (TEDDY) Pub92 Agardh

Prepared by Allyson Mateja
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD 20705
February 22, 2017

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 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.

3 Archived Datasets

All the 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_92_dagardh_niddk_31july2014_1.sas7bdat" and "m_92_dagaedh_niddk_31july2014_2.sas7bdat" datasets.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Agardh et al [1] in Pediatrics in 2015. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Clinical Presentation at Time of Seroconversion in Children Persistently Positive for tTGA in 2 Consecutive Samples (i.e., CDA) Compared With Children Negative for tTGA (No CDA), Table A lists the variables that were used in the replication and Table B compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are almost an exact match to the published results.

For Table 2 in the publication [1], Association Between Identified Risk Factors for CD and Having ≥ 1 Symptoms at Time of Seroconversion in Children With CDA Respective in CDA Children Who Were Diagnosed with CD, Table C lists the variables that were used in the replication and Table D compares the results calculated from the archived data file to the results published in Table 2. The results of the replication are an exact match to the published results.

For Table 3 in the publication [1], Paired Comparison of the Clinical Presentation in Children With CDA Who Had Completed the Symptom Questionnaires at 3 Time Points (n = 205), Table E lists the variables that were used in the replication and Table F compares the results calculated from the archived data file to the results published in Table 3. The results of the replication are almost an exact match to the published results.

For Table 4 in the publication [1], Factors Associated With Levels of tTGA at Time of Seroconversion in Children With CDA and in Children With CDA Who Were Diagnosed With CD, Table G lists the variables that were used in the replication and Table H compares the results calculated from the archived data file to the results published in Table 4. The results of the replication are almost an exact match to the published results.

6 Conclusions

The NIDDK repository is confident that the TEDDY M92 data files to be distributed are a true copy of the study data.

7 References

[1] Agardh, D., Lee, H., Kurppa, K., Simell, V., Aronsson, C.A., Jorneus, O., Hummel, M., Liu, E., Koletzko, S., and the TEDDY study group. "Clinical Features of Celiac Disease: A Prospective Birth Cohort". Pediatrics (2015) 135(4):627-634.

Table A: Variables used to replicate Table 1**:** Clinical Presentation at Time of Seroconversion in Children Persistently Positive for tTGA in 2 Consecutive Samples (i.e., CDA) Compared With Children Negative for tTGA (No CDA)

Table Variable	**dataset.variable**
By 2 y of Age, By 3 y of Age, By 4 y of Age	m_92_dagaedh_niddk_31july2014_2.due_num
CDA	m_92_dagaedh_niddk_31july2014_1.celiac_disease, m_92_dagaedh_niddk_31july2014_1.mtimecd
No CDA	m_92_dagaedh_niddk_31july2014_1.celiac_disease, m_92_dagaedh_niddk_31july2014_1.persist_tga, m_92_dagaedh_niddk_31july2014_1.init_pos
Wt z-score	m_92_dagaedh_niddk_31july2014_2.waz
Height, z-score	m_92_dagaedh_niddk_31july2014_2.haz
BMI, z-score	m_92_dagaedh_niddk_31july2014_2.bmiz
≥ 1 Symptoms	m_92_dagaedh_niddk_31july2014_2.nosymptoms
Abdominal pain	m_92_dagaedh_niddk_31july2014_2.abdominaldiscomfort
Anemia	m_92_dagaedh_niddk_31july2014_2.anemia
Constipation	m_92_dagaedh_niddk_31july2014_2.chronicconstipation
Dental enamel defects	m_92_dagaedh_niddk_31july2014_2.dentalenemaldefects
Fatigue	m_92_dagaedh_niddk_31july2014_2.fatigue
Loose stools	m_92_dagaedh_niddk_31july2014_2.frequentloosestools
Irritability	m_92_dagaedh_niddk_31july2014_2.irritability
Neurological symptoms	m_92_dagaedh_niddk_31july2014_2.neurologicalsymptoms
Poor growth	m_92_dagaedh_niddk_31july2014_2.poorgrowth
Skin irritation	m_92_dagaedh_niddk_31july2014_2.skinirritation
Vomiting	m_92_dagaedh_niddk_31july2014_2.vomiting
Other	m_92_dagaedh_niddk_31july2014_2.other

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

By 2 y of Age

Variable	CDA Manuscript, n=115	CDA DSIC, n=115	Diff. n=0	No CDA Manuscript, n=1932	No CDA DSIC, n=1932	Diff. n=0
Growth, mean z-score (SD)						
Wt	0.5 (1.1)	0.5 (1.1)	0 (0)	0.2 (1.0)	0.2 (1.0)	0 (0)
Height	0.4 (0.8)	0.4 (0.8)	0 (0)	0.4 (0.9)	0.4 (0.9)	0 (0)
BMI	0.3 (1.1)	0.3 (1.1)	0 (0)	0.1 (1.0)	0.1 (1.0)	0 (0)
Growth < 10th percentile, %						
Wt	5	5	0	6	6	0
Height	3	3	0	3	3	0

Variable	CDA Manuscript, n=115	CDA DSIC, n=115	Diff. n=0	No CDA Manuscript, n=1932	No CDA DSIC, n=1932	Diff. n=0
BMI	7	7	0	6	6	0
Clinical presentation, %						
≥ 1 Symptoms	34	34	0	19	19	0
Abdominal pain	10	10	0	5	5	0
Anemia	2	2	0	1	1	0
Constipation	11	11	0	6	6	0
Dental enamel defects	1	1	0	1	1	0
Fatigue	1	1	0	1	1	0
Loose stools	13	13	0	7	7	0
Irritability	3	3	0	2	2	0
Neurological symptoms	0	0	0	0.4	0.4	0
Poor growth	6	6	0	2	2	0
Skin irritation	3	3	0	2	2	0
Vomiting	3	3	0	1	1	0
Other	1	1	0	0.3	0.3	0

By 3 y of Age

Variable	CDA Manuscript, n=142	CDA DSIC, n=142	Diff. n=0	No CDA Manuscript, n=2794	No CDA DSIC, n=2794	Diff. n=0
Growth, mean z-score (SD)						
Wt	0.4 (0.9)	0.4 (0.9)	0 (0)	0.5 (1.0)	0.5 (1.0)	0 (0)
Height	0.4 (1.0)	0.4 (1.0)	0 (0)	0.4 (0.9)	0.4 (0.9)	0 (0)
BMI	0.2 (0.8)	0.2 (0.8)	0 (0)	0.2 (1.0)	0.2 (1.0)	0 (0)
Growth < 10th percentile, %						
Wt	6	6	0	4	4	0
Height	6	6	0	3	3	0
BMI	4	4	0	7	6	1
Clinical presentation, %						
≥ 1 Symptoms	28	28	0	19	19	0
Abdominal pain	16	16	0	8	8	0
Anemia	1	1	0	1	1	0
Constipation	6	6	0	6	6	0
Dental enamel defects	2	2	0	1	1	0
Fatigue	2	2	0	1	1	0
Loose stools	6	6	0	3	3	0
Irritability	4	4	0	2	2	0

Variable	CDA Manuscript, n=142	CDA DSIC, n=142	Diff. n=0	No CDA Manuscript, n=2794	No CDA DSIC, n=2794	Diff. n=0
Neurological symptoms	1	1	0	0.3	0.3	0
Poor growth	8	8	0	2	2	0
Skin irritation	1	1	0	1	1	0
Vomiting	1	1	0	1	1	0
Other	1	1	0	0.1	0.1	0

By 4 y of Age

Variable	CDA Manuscript, n=131	CDA DSIC, n=131	Diff. n=0	No CDA Manuscript, n=3355	No CDA DSIC, n=3355	Diff. n=0
Growth, mean z-score (SD)						
Wt	0.4 (1.0)	0.4 (1.0)	0 (0)	0.5 (0.9)	0.5 (0.9)	0 (0)
Height	0.6 (1.0)	0.6 (1.0)	0 (0)	0.5 (1.0)	0.5 (1.0)	0 (0)
BMI	0.2 (1.0)	0.2 (1.0)	0 (0)	0.2 (1.0)	0.2 (1.0)	0 (0)
Growth < 10th percentile, %						
Wt	4	4	0	3	3	0
Height	4	4	0	3	3	0
BMI	6	6	0	6	6	0
Clinical presentation, %						
≥ 1 Symptoms	27	27	0	21	21	0
Abdominal pain	15	15	0	11	11	0
Anemia	0	0	0	1	1	0
Constipation	8	8	0	5	5	0
Dental enamel defects	2	2	0	2	2	0
Fatigue	0	0	0	2	2	0
Loose stools	4	4	0	3	3	0
Irritability	5	5	0	3	3	0
Neurological symptoms	1	1	0	0.2	0.2	0
Poor growth	3	3	0	1	1	0
Skin irritation	2	2	0	1	1	0
Vomiting	3	3	0	27 (1)	27 (1)	0 (0)
Other	0	0	0	2 (0.1)	2 (0.1)	0 (0)

Table C: Variables used to replicate Table 2**:** Association Between Identified Risk Factors for CD and Having ≥ 1 Symptoms at Time of Seroconversion in Children With CDA Respective in CDA Children Who Were Diagnosed with CD

Table Variable	**dataset.variable**
CDA	m_92_dagaedh_niddk_31july2014_1.persist_visit1
CD	m_92_dagaedh_niddk_31july2014_1.celiac_disease, m_92_dagaedh_niddk_31july2014_1.marsh
Country	m_92_dagaedh_niddk_31july2014_1.country
Gender	m_92_dagaedh_niddk_31july2014_1.sex
HLA-DR3-DQ2	m_92_dagaedh_niddk_31july2014_1.hlarg
FDR with CD	m_92_dagaedh_niddk_31july2014_1.celiac_fdr
≥ 1 Symptoms	m_92_dagaedh_niddk_31july2014_2.nosymptoms

Table D: Comparison of values computed in integrity check to reference article Table 2 values

CDA Children

Variable		n Manuscript	n DSIC	Diff.	≥ 1 Symptoms Manuscript	≥ 1 Symptoms DSIC	Diff.
Country	US	170	170	0	55 (32)	55 (32)	0 (0)
	Finland	96	96	0	29 (30)	29 (30)	0 (0)
	Germany	21	21	0	4 (19)	4 (19)	0 (0)
	Sweden	171	171	0	28 (16)	28 (16)	0 (0)
Gender	Male	194	194	0	47 (24)	47 (24)	0 (0)
	Female	264	264	0	69 (26)	69 (26)	0 (0)
HLA-DR3-DQ2 (n)	0	104	104	0	22 (21)	22 (21)	0 (0)
	1	176	176	0	53 (30)	53 (30)	0 (0)
	2	178	178	0	41 (23)	41 (23)	0 (0)
FDR with CD	No	423	423	0	101 (24)	101 (24)	0 (0)
	Yes	29	29	0	13 (45)	13 (45)	0 (0)

		OR (95% CI) Manuscript	OR (95% CI) DSIC	Diff.
Country	US	Reference	Reference	NA
	Finland	0.91 (0.53-1.56)	0.91 (0.53-1.56)	0 (0-0)
	Germany	0.49 (0.16-1.53)	0.49 (0.16-1.53)	0 (0-0)
	Sweden	0.41 (0.24-0.69)	0.41 (0.24-0.69)	0 (0-0)
Gender	Male	Reference	Reference	NA
	Female	1.11 (0.72-1.70)	1.11 (0.72-1.70)	0 (0-0)
HLA-DR3-DQ2 (n)	0	Reference	Reference	NA
	1	1.61 (0.91-2.84)	1.61 (0.91-2.84)	0 (0-0)

		OR (95% CI) Manuscript	OR (95% CI) DSIC	Diff.
	2	1.12 (0.62-2.00)	1.12 (0.62-2.00)	0 (0-0)
FDR with CD	No	Reference	Reference	NA
	Yes	2.59 (1.21-5.57)	2.59 (1.21-5.57)	0 (0-0)

CD Children

Variable		n Manuscript	n DSIC	Diff.	≥ 1 Symptoms Manuscript	≥ 1 Symptoms DSIC	Diff.
Country	US	52	52	0	24 (46)	24 (46)	0 (0)
	Finland	28	28	0	8 (29)	8 (29)	0 (0)
	Germany	3	3	0	0 (0)	0 (0)	0 (0)
	Sweden	72	72	0	10 (14)	10 (14)	0 (0)
Gender	Male	63	63	0	16 (25)	16 (25)	0 (0)
	Female	92	92	0	26 (28)	26 (28)	0 (0)
HLA-DR3-DQ2 (n)	0	24	24	0	6 (25)	6 (25)	0 (0)
	1	60	60	0	19 (32)	19 (32)	0 (0)
	2	71	71	0	17 (24)	17 (24)	0 (0)
FDR with CD	No	138	138	0	33 (24)	33 (24)	0 (0)
	Yes	16	16	0	8 (50)	8 (50)	0 (0)

		OR (95% CI) Manuscript	OR (95% CI) DSIC	Diff.
Country	US	Reference	Reference	NA
	Finland	0.47 (0.17-1.25)	0.47 (0.17-1.25)	0 (0-0)
	Germany	NA	NA	NA
	Sweden	0.19 (0.08-0.45)	0.19 (0.08-0.45)	0 (0-0)
Gender	Male	Reference	Reference	NA
	Female	1.16 (0.56-2.39)	1.16 (0.56-2.39)	0 (0-0)
HLA-DR3-DQ2 (n)	0	Reference	Reference	NA
	1	1.39 (0.48-4.06)	1.39 (0.48-4.06)	0 (0-0)
	2	0.94 (0.32-2.76)	0.94 (0.32-2.76)	0 (0-0)
FDR with CD	No	Reference	Reference	NA
	Yes	3.18 (1.11-9.14)	3.18 (1.11-9.14)	0 (0-0)

Table E: Variables used to replicate Table 3**:** Paired Comparison of the Clinical Presentation in Children With CDA Who Had Completed the Symptom Questionnaires at 3 Time Points (n = 205)

Table Variable	**dataset.variable**
Time of Initial tTGA Positivity	m_92_dagaedh_niddk_31july2014_1.persist_visit1
Time of Persistent tTGA Positivity	m_92_dagaedh_niddk_31july2014_1.persist_visit2
Wt z-score	m_92_dagaedh_niddk_31july2014_2.waz
Height, z-score	m_92_dagaedh_niddk_31july2014_2.haz
BMI, z-score	m_92_dagaedh_niddk_31july2014_2.bmiz
≥ 1 Symptoms	m_92_dagaedh_niddk_31july2014_2.nosymptoms
Abdominal pain	m_92_dagaedh_niddk_31july2014_2.abdominaldiscomfort
≥ 1 Symptoms	m_92_dagaedh_niddk_31july2014_2.nosymptoms
Abdominal pain	m_92_dagaedh_niddk_31july2014_2.abdominaldiscomfort
Anemia	m_92_dagaedh_niddk_31july2014_2.anemia
Constipation	m_92_dagaedh_niddk_31july2014_2.chronicconstipation
Dental enamel defects	m_92_dagaedh_niddk_31july2014_2.dentalenemaldefects
Fatigue	m_92_dagaedh_niddk_31july2014_2.fatigue
Loose stools	m_92_dagaedh_niddk_31july2014_2.frequentloosestools
Irritability	m_92_dagaedh_niddk_31july2014_2.irritability
Neurological symptoms	m_92_dagaedh_niddk_31july2014_2.neurologicalsymptoms
Poor growth	m_92_dagaedh_niddk_31july2014_2.poorgrowth
Skin irritation	m_92_dagaedh_niddk_31july2014_2.skinirritation
Vomiting	m_92_dagaedh_niddk_31july2014_2.vomiting
Other	m_92_dagaedh_niddk_31july2014_2.other

Table F: Comparison of values computed in integrity check to reference article Table 3 values

Clinical Presentation	At 12 mo Before Initial tTGA Positivity, n (%) Manuscript	At 12 mo Before Initial tTGA Positivity, n (%) DSIC	Diff.	At Time of Initial tTGA Positivity, n (%) Manuscript	At Time of Initial tTGA Positivity, n (%) DSIC	Diff.
Wt, z-score (SD)	0.38 (0.90)	0.38 (0.90)	0 (0)	0.41 (0.85)	0.41 (0.85)	0 (0)
Height, z-score (SD)	0.56 (0.90)	0.56 (0.90)	0 (0)	0.57 (0.91)	0.57 (0.91)	0 (0)
BMI, z-score (SD)	0.06 (0.97)	0.06 (0.97)	0 (0)	0.12 (0.90)	0.12 (0.90)	0 (0)
≥ 1 Symptoms	41 (20)	42 (20)	1 (0)	43 (21)	43 (21)	0 (0)
Abdominal discomfort	17 (8)	17 (8)	0 (0)	26 (12)	26 (13)	0 (1)
Anemia	1 (0.5)	1 (0.5)	0 (0)	2 (1)	2 (1)	0 (0)
Constipation	12 (6)	12 (6)	0 (0)	14 (7)	14 (7)	0 (0)
Dental enamel defects	1 (0.5)	1 (0.5)	0 (0)	3 (1)	3 (1)	0 (0)
Fatigue	4 (2)	4 (2)	0 (0)	3 (1)	3 (1)	0 (0)
Frequent loose stools	12 (6)	12 (6)	0 (0)	6 (3)	6 (3)	0 (0)
Irritability	4 (2)	4 (2)	0 (0)	8 (4)	8 (4)	0 (0)

Clinical Presentation	At 12 mo Before Initial tTGA Positivity, n (%) Manuscript	At 12 mo Before Initial tTGA Positivity, n (%) DSIC	Diff.	At Time of Initial tTGA Positivity, n (%) Manuscript	At Time of Initial tTGA Positivity, n (%) DSIC	Diff.
Neurological symptoms	2 (1)	2 (1)	0 (0)	1 (0.5)	1 (0.5)	0 (0)
Poor growth	1 (0.5)	1 (0.5)	0 (0)	2 (1)	2 (1)	0 (0)
Skin irritation	3 (1)	3 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Vomiting	3 (1)	3 (1)	0 (0)	3 (1)	3 (1)	0 (0)
Other	1 (0.5)	1 (0.5)	0 (0)	0 (0)	0 (0)	0 (0)

Clinical Presentation	At Time of Persistent tTGA Positivity, n (%) Manuscript	At Time of Persistent tTGA Positivity, n (%) DSIC	Diff.
Wt, z-score (SD)	0.40 (0.83)	0.40 (0.83)	0 (0)
Height, z-score (SD)	0.56 (0.89)	0.56 (0.89)	0 (0)
BMI, z-score (SD)	0.14 (0.90)	0.14 (0.90)	0 (0)
≥ 1 Symptoms	88 (43)	88 (43)	0 (0)
Abdominal discomfort	59 (29)	59 (29)	0 (0)
Anemia	1 (0.5)	1 (0.5)	0 (0)
Constipation	22 (11)	22 (11)	0 (0)
Dental enamel defects	5 (2)	5 (2)	0 (0)
Fatigue	8 (4)	8 (4)	0 (0)
Frequent loose stools	23 (11)	23 (11)	0 (0)
Irritability	15 (7)	15 (7)	0 (0)
Neurological symptoms	1 (0.5)	1 (0.5)	0 (0)
Poor growth	3 (1)	3 (1)	0 (0)
Skin irritation	0 (0)	0 (0)	0 (0)
Vomiting	3 (1)	3 (1)	0 (0)
Other	1 (0.5)	1 (0.5)	0 (0)

Table G: Variables used to replicate Table 4**:** Factors Associated With Levels of tTGA at Time of Seroconversion in Children With CDA and in Children With CDA Who Were Diagnosed With CD

Table Variable	**dataset.variable**
Wt < 10^{th} percentile	m_92_dagaedh_niddk_31july2014_2.waz
Height < 10^{th} percentile	m_92_dagaedh_niddk_31july2014_2.haz
BMI < 10^{th} percentile	m_92_dagaedh_niddk_31july2014_2.bmiz
≥ 1 Symptoms	m_92_dagaedh_niddk_31july2014_2.nosymptoms
Country	m_92_dagaedh_niddk_31july2014_1.country
Gender	m_92_dagaedh_niddk_31july2014_1.sex

Table Variable	dataset.variable
HLA-DR3-DQ2	m_92_dagaedh_niddk_31july2014_1.hlarg
FDR with CD	m_92_dagaedh_niddk_31july2014_1.celiac_fdr
tTGA level	m_92_dagaedh_niddk_31july2014_1.persist_bristol1
Celiac Disease	m_92_dagaedh_niddk_31july2014_1.celiac_disease, m_92_dagaedh_niddk_31july2014_1.marsh
CDA	m_92_dagaedh_niddk_31july2014_1.persist_visit1

Table H: Comparison of values computed in integrity check to reference article Table 4 values

Children With CDA

Variable		n Manuscript	n DSIC	Diff.	tTGA Level Median (Q1, Q3) Manuscript	tTGA Level Median (Q1, Q3) DSIC	Diff.
Wt < 10th percentile	No	428	428	0	16.5 (5.1, 48.7)	16.5 (5.1, 48.7)	0 (0, 0)
	Yes	17	17	0	35.8 (16.5, 78.7)	35.8 (16.5, 78.7)	0 (0, 0)
Height < 10th percentile	No	426	426	0	16.3 (5, 49.8)	16.3 (5.0, 49.8)	0 (0, 0)
	Yes	15	15	0	32.6 (20.8, 94.6)	32.6 (20.8, 94.6)	0 (0, 0)
BMI < 10th percentile	No	396	396	0	16.5 (4.9, 49.6)	16.5 (4.9, 49.6)	0 (0, 0)
	Yes	20	20	0	31.6 (17.0, 72.8)	31.5 (17.0, 72.8)	0.1 (0, 0)
≥ 1 Symptoms	No	342	342	0	14.9 (4.9, 41.1)	14.9 (4.9, 41.1)	0 (0, 0)
	Yes	116	116	0	29.1 (6.4, 79.1)	29.1 (6.4, 79.1)	0 (0, 0)
Country	US	170	170	0	30.0 (11.4, 70.3)	30.0 (11.4, 70.3)	0 (0, 0)
	Finland	96	96	0	10.1 (3.9, 31.7)	10.1 (3.9, 31.7)	0 (0, 0)
	Germany	21	21	0	13.6 (4.0, 29.5)	13.6 (4.0, 29.5)	0 (0, 0)
	Sweden	171	171	0	13.9 (4.3, 41.0)	13.9 (4.3, 41.0)	0 (0, 0)
Gender	Male	194	194	0	16.5 (5.4, 41.4)	16.5 (5.4, 41.1)	0 (0, 0.3)
	Female	264	264	0	18.9 (5.0, 57.4)	18.9 (5.0, 57.4)	0 (0, 0)
HLA-DR3-DQ2 (n)	0	104	104	0	18.4 (4.7, 39.9)	18.4 (4.7, 39.9)	0 (0, 0)
	1	176	176	0	17.6 (5.2, 54.4)	17.5 (5.2, 54.4)	0.1 (0, 0)
	2	178	178	0	17.4 (5.3, 55.2)	17.4 (5.3, 55.2)	0 (0, 0)
FDR with CD	No*	423	423	0	6.2 (5.0, 47.2)	16.2 (5.0, 47.2)	10.0 (0, 0)
	Yes	29	29	0	49.8 (20.2, 97.5)	49.8 (20.2, 97.5)	0 (0, 0)

Children With CD

Variable		n Manuscript	n DSIC	Diff.	tTGA Level Median (Q1, Q3) Manuscript	tTGA Level Median (Q1, Q3) DSIC	Diff.
Wt < 10th percentile	No	146	146	0	43.6 (14.9, 91.3)	43.6 (14.9, 91.3)	0 (0, 0)
	Yes	6	6	0	63.8 (20.8, 100.5)	63.7 (20.8, 100.5)	0.1 (0, 0)
Height < 10th percentile	No	140	140	0	45.5 (14.7, 86.7)	45.5 (14.7, 86.7)	0 (0, 0)
	Yes	9	9	0	48.2 (21.4, 100.5)	48.2 (21.3, 100.5)	0 (0.1, 0)
BMI < 10th percentile	No	132	132	0	45.5 (15.1, 91.7)	45.5 (15.1, 91.7)	0 (0, 0)
	Yes*	7	7	0	82.3 (57.5, 137.8)	87.3 (57.5, 137.8)	5 (0, 0)
≥ 1 Symptoms	No	113	113	0	39.1 (11.5, 74.2)	39.1 (11.5, 74.2)	0 (0, 0)
	Yes	42	42	0	81.5 (32.1, 135.4)	81.5 (32.1, 135.4)	0 (0, 0)
Country	US	52	52	0	63.0 (31.6, 116.7)	63.0 (31.6, 116.7)	0 (0, 0)
	Finland	28	28	0	29.4 (9.6, 81.4)	29.4 (9.6, 81.4)	0 (0, 0)
	Germany	3	3	0	78.0 (75.2, 98.4)	78.0 (75.2, 98.4)	0 (0, 0)
	Sweden	72	72	0	34.7 (9.7, 70.5)	34.7 (9.7, 70.5)	0 (0, 0)
Gender	Male	63	63	0	38.9 (13.9, 100.5)	38.9 (13.9, 100.5)	0 (0, 0)
	Female	92	92	0	51.0 (16.7, 78.6)	51.0 (16.7, 78.6)	0 (0, 0)
HLA-DR3-DQ2 (n)	0	24	24	0	70.2 (22.3, 101.3)	70.1 (22.3, 101.3)	0.1 (0, 0)
	1	60	60	0	52.4 (20.7, 93.4)	52.4 (20.7, 93.4)	0 (0, 0)
	2	71	71	0	34.0 (7.9, 78.0)	34.0 (7.9, 78.0)	0 (0, 0)
FDR with CD	No	138	138	0	42.0 (14.9, 86.1)	42.0 (14.9, 86.1)	0 (0, 0)
	Yes	16	16	0	89.1 (37.3, 118.4)	89.1 (37.3, 118.4)	0 (0, 0)

*The DCC confirmed that the numbers presented in the manuscript are incorrect. The numbers calculated in the DSIC are the correct values.

Attachment A: SAS Code

```
*** TEDDY M92 DSIC;
*** Programmer: Allyson Mateja;
*** Date: 11/7/16;
*** Modified: 1/4/2017;

proc format;
       value countryf 1 = 'US'
                      2 = 'Finland'
                      3 = 'Germany'
                      4 = 'Sweden';

       value hlaf 0,2,3,5,6,7,9,10  = 0
                  1                 = 1
                  4                 = 2;

libname m92data '/prj/niddk/ims_analysis/TEDDY/private_orig_data/m_92_dagardh_niddk_submission';

data m92_data1;
       set m92data.m_92_dagardh_niddk_31july2014_1;

data m92_data2;
       set m92data.m_92_dagardh_niddk_31july2014_2;

proc contents data = m92_data1;
proc contents data = m92_data2;

proc freq data = m92_data2;
       tables due_num;

proc freq data = m92_data1;
       tables persist_tga persist_tga*marsh*celiac_disease /list missing;
       title 'Table 1';

proc sort data = m92_data2 nodupkey out = m92_data2_ids;
       by maskid;

data m92_data1;
       set m92_data1;
       twelve_mo_prior = persist_visit1 - 12;

proc sort data=m92_data1;
       by maskid persist_visit1;

proc sort data=m92_data2 nodupkey;
       by maskid due_num;

data table1;
```

```
        merge m92_data1 (in=val1 rename = (persist_visit1 = due_num))
              m92_data2 (in=val2);
        by maskid due_num;
        num_symptoms = sum(abdominaldiscomfort, anemia, chronicconstipation, dentalenemaldefects, fatigue, frequentloosestools,
irritability, neurologicalsymptoms, poorgrowth, skinirritation, vomiting, other);
        if nosymptoms = . then ge1_symptom = 1;
        else if nosymptoms = 1 then ge1_symptom = 0;
        if hlarg = . then hlarg = 0;
        if . < waz <= -1.28 then wt_less_10_per = 1;
        else if waz > -1.28 then wt_less_10_per = 0;
        if . < bmiz <= -1.28 then bmi_less_10_per = 1;
        else if bmiz > -1.28 then bmi_less_10_per = 0;
        if . < haz <= -1.28 then ht_less_10_per = 1;
        else if haz > -1.28 then ht_less_10_per = 0;
        if val1 and val2 and due_num ne . then output table1;

data determine_ages;
        merge m92_data2 (in=val1)
              m92_data1 (in=val2);
        by maskid;
        if nosymptoms = . then ge1_symptom = 1;
        else if nosymptoms = 1 then ge1_symptom = 0;
        if . < waz <= -1.28 then wt_less_10_per = 1;
        else if waz > -1.28 then wt_less_10_per = 0;
        if . < bmiz <= -1.28 then bmi_less_10_per = 1;
        else if bmiz > -1.28 then bmi_less_10_per = 0;
        if . < haz <= -1.28 then ht_less_10_per = 1;
        else if haz > -1.28 then ht_less_10_per = 0;
        if val1 and val2 and cc ne 134 then output;

proc sort data=determine_ages nodupkey;
        by maskid due_num;

data by_2yrage by_3yrage by_4yrage;
        set determine_ages;
        if due_num = 24 then output by_2yrage;
        else if due_num=36 then output by_3yrage;
        else if due_num=48 then output by_4yrage;

data by_2yrage;
        set by_2yrage;
        if (persist_visit1^=. and persist_visit1<=24) then case24=1;
        else if (celiac_disease=1 and mtimecd<=26) then case24=1;
        else if (persist_tga^=1 or init_pos^=1 or celiac_disease^=1) then case24=0;

data by_3yrage;
        set by_3yrage;
        if (persist_visit1^=. and persist_visit1<=24) then delete;
        else if (celiac_disease=1 and mtimecd<=26) then delete;
        else if (persist_visit1^=. and 24<persist_visit1=36) then case36=1;
        else if (celiac_disease=1 and 26<mtimecd<=38) then case36=1;
```

```
        else if (persist_tga^=1 or init_pos^=1 or celiac_disease^=1) then case36=0;

data by_4yrage;
        set by_4yrage;
        if (persist_visit1^=. and persist_visit1<=24) then delete;
        else if (celiac_disease=1 and mtimecd<=26) then delete;
        else if (persist_visit1^=. and 24<persist_visit1=36) then delete;
        else if (celiac_disease=1 and 26<mtimecd<=38) then delete;
        else if (persist_visit1^=. and 36<persist_visit1<=48) then case48=1;
        else if (celiac_disease=1 and 38<mtimecd<=50) then case48=1;
        else if (persist_tga^=1 or init_pos^=1 or celiac_disease^=1) then case48=0;

proc freq data = by_2yrage;
        tables case24;
        title 'Table 1 - By 2 y of Age';

proc freq data = by_3yrage;
        tables case36;
        title 'Table 1 - By 3 y of Age';

proc freq data = by_4yrage;
        tables case48;
        title 'Table 1 - By 4 y of Age';

proc means data = by_2yrage n mean std;
        var waz;
        class case24;
        where case24 ne .;
        title 'Table 1 - Wt By 2 y of Age';

proc means data = by_2yrage n mean std;
        var haz;
        class case24;
        where case24 ne .;
        title 'Table 1 - Height By 2 y of Age';

proc means data = by_2yrage n mean std;
        var bmiz;
        class case24;
        where case24 ne .;
        title 'Table 1 - BMI By 2 y of Age';

proc means data = by_3yrage n mean std;
        var waz;
        class case36;
        where case36 ne .;
        title 'Table 1 - Wt By 3 y of Age';

proc means data = by_3yrage n mean std;
        var haz;
        class case36;
```

```
        where case36 ne .;
        title 'Table 1 - Height By 3 y of Age';

proc means data = by_3yrage n mean std;
        var bmiz;
        class case36;
        where case36 ne .;
        title 'Table 1 - BMI By 3 y of Age';

proc means data = by_4yrage n mean std;
        var waz;
        class case48;
        where case48 ne .;
        title 'Table 1 - Wt By 4 y of Age';

proc means data = by_4yrage n mean std;
        var haz;
        class case48;
        where case48 ne .;
        title 'Table 1 - Height By 4 y of Age';

proc means data = by_4yrage n mean std;
        var bmiz;
        class case48;
        where case48 ne .;
        title 'Table 1 - BMI By 4 y of Age';

proc sort data = by_2yrage;
        by case24;

proc freq data = by_2yrage;
        tables wt_less_10_per /missing;
        by case24;
        where case24 ne .;
        title 'Table 1 - Wt < 10th percentile, By 2 yr of Age';

proc freq data = by_2yrage;
        tables ht_less_10_per /missing;
        by case24;
        where case24 ne .;
        title 'Table 1 - Height < 10th percentile, By 2 yr of Age';

proc freq data = by_2yrage;
        tables bmi_less_10_per /missing;
        by case24;
        where case24 ne .;
        title 'Table 1 - BMI < 10th percentile, By 2 yr of Age';

proc sort data = by_3yrage;
        by case36;
```

```
proc freq data = by_3yrage;
       tables wt_less_10_per /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Wt < 10th percentile, By 3 yr of Age';

proc freq data = by_3yrage;
       tables ht_less_10_per /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Height < 10th percentile, By 3 yr of Age';

proc freq data = by_3yrage;
       tables bmi_less_10_per /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - BMI < 10th percentile, By 3 yr of Age';

proc sort data = by_4yrage;
       by case48;

proc freq data = by_4yrage;
       tables wt_less_10_per /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - Wt < 10th percentile, By 4 yr of Age';

proc freq data = by_4yrage;
       tables ht_less_10_per /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - Height < 10th percentile, By 4 yr of Age';

proc freq data = by_4yrage;
       tables bmi_less_10_per /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - BMI < 10th percentile, By 4 yr of Age';

proc freq data = by_2yrage;
       tables ge1_symptom /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - >= 1 Symptoms, By 2 yr of Age';

proc freq data = by_2yrage;
       tables abdominaldiscomfort /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Abdominal pain, By 2 yr of Age';
```

```
proc freq data = by_2yrage;
       tables anemia /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Anemia, By 2 yr of Age';

proc freq data = by_2yrage;
       tables chronicconstipation /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Constipation, By 2 yr of Age';

proc freq data = by_2yrage;
       tables dentalenemaldefects /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Dental enamel defects, By 2 yr of Age';

proc freq data = by_2yrage;
       tables fatigue /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Fatigue, By 2 yr of Age';

proc freq data = by_2yrage;
       tables frequentloosestools /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Loose stools, By 2 yr of Age';

proc freq data = by_2yrage;
       tables irritability /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Irritability, By 2 yr of Age';

proc freq data = by_2yrage;
       tables neurologicalsymptoms /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Neurological symptoms, By 2 yr of Age';

proc freq data = by_2yrage;
       tables poorgrowth /missing;
       by case24;
       where case24 ne .;
       title 'Table 1 - Poor growth, By 2 yr of Age';

proc freq data = by_2yrage;
       tables skinirritation /missing;
       by case24;
```

```
        where case24 ne .;
        title 'Table 1 - Skin irritation, By 2 yr of Age';

proc freq data = by_2yrage;
        tables vomiting /missing;
        by case24;
        where case24 ne .;
        title 'Table 1 - Vomiting, By 2 yr of Age';

proc freq data = by_2yrage;
        tables other /missing;
        by case24;
        where case24 ne .;
        title 'Table 1 - Other, By 2 yr of Age';

proc freq data = by_3yrage;
        tables ge1_symptom /missing;
        by case36;
        where case36 ne .;
        title 'Table 1 - >= 1 Symptoms, By 3 yr of Age';

proc freq data = by_3yrage;
        tables abdominaldiscomfort /missing;
        by case36;
        where case36 ne .;
        title 'Table 1 - Abdominal pain, By 3 yr of Age';

proc freq data = by_3yrage;
        tables anemia /missing;
        by case36;
        where case36 ne .;
        title 'Table 1 - Anemia, By 3 yr of Age';

proc freq data = by_3yrage;
        tables chronicconstipation /missing;
        by case36;
        where case36 ne .;
        title 'Table 1 - Constipation, By 3 yr of Age';

proc freq data = by_3yrage;
        tables dentalenemaldefects /missing;
        by case36;
        where case36 ne .;
        title 'Table 1 - Dental enamel defects, By 3 yr of Age';

proc freq data = by_3yrage;
        tables fatigue /missing;
        by case36;
        where case36 ne .;
        title 'Table 1 - Fatigue, By 3 yr of Age';
```

```
proc freq data = by_3yrage;
       tables frequentloosestools /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Loose stools, By 3 yr of Age';

proc freq data = by_3yrage;
       tables irritability /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Irritability, By 3 yr of Age';

proc freq data = by_3yrage;
       tables neurologicalsymptoms /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Neurological symptoms, By 3 yr of Age';

proc freq data = by_3yrage;
       tables poorgrowth /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Poor growth, By 3 yr of Age';

proc freq data = by_3yrage;
       tables skinirritation /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Skin irritation, By 3 yr of Age';

proc freq data = by_3yrage;
       tables vomiting /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Vomiting, By 3 yr of Age';

proc freq data = by_3yrage;
       tables other /missing;
       by case36;
       where case36 ne .;
       title 'Table 1 - Other, By 3 yr of Age';

proc freq data = by_4yrage;
       tables ge1_symptom /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - >= 1 Symptoms, By 4 yr of Age';

proc freq data = by_4yrage;
       tables abdominaldiscomfort /missing;
       by case48;
```

```
        where case48 ne .;
        title 'Table 1 - Abdominal pain, By 4 yr of Age';

proc freq data = by_4yrage;
        tables anemia /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Anemia, By 4 yr of Age';

proc freq data = by_4yrage;
        tables chronicconstipation /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Constipation, By 4 yr of Age';

proc freq data = by_4yrage;
        tables dentalenemaldefects /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Dental enamel defects, By 4 yr of Age';

proc freq data = by_4yrage;
        tables fatigue /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Fatigue, By 4 yr of Age';

proc freq data = by_4yrage;
        tables frequentloosestools /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Loose stools, By 4 yr of Age';

proc freq data = by_4yrage;
        tables irritability /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Irritability, By 4 yr of Age';

proc freq data = by_4yrage;
        tables neurologicalsymptoms /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Neurological symptoms, By 4 yr of Age';

proc freq data = by_4yrage;
        tables poorgrowth /missing;
        by case48;
        where case48 ne .;
        title 'Table 1 - Poor growth, By 4 yr of Age';
```

```
proc freq data = by_4yrage;
       tables skinirritation /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - Skin irritation, By 4 yr of Age';

proc freq data = by_4yrage;
       tables vomiting /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - Vomiting, By 4 yr of Age';

proc freq data = by_4yrage;
       tables other /missing;
       by case48;
       where case48 ne .;
       title 'Table 1 - Other, By 4 yr of Age';

proc sort data=m92_data1;
       by maskid twelve_mo_prior;

data table3_part1;
       merge m92_data1 (in=val1 rename = (twelve_mo_prior = due_num))
             m92_data2 (in=val2);
       by maskid due_num;
       if nosymptoms = . then ge1_symptom = 1;
       else if nosymptoms = 1 then ge1_symptom = 0;
       if val1 and val2 and due_num ne . and cc ne 134 then output;

proc sort data=m92_data1;
       by maskid persist_visit2;

data table3_part2;
       merge m92_data1 (in=val1 rename = (persist_visit2 = due_num))
             m92_data2 (in=val2);
       by maskid due_num;
       if nosymptoms = . then ge1_symptom = 1;
       else if nosymptoms = 1 then ge1_symptom = 0;
       if val1 and val2 and due_num ne . and cc ne 134 then output;

data all_three_visits;
       merge table1       (in=val1 keep=maskid waz haz bmiz ge1_symptom abdominaldiscomfort anemia chronicconstipation
dentalenemaldefects fatigue frequentloosestools irritability neurologicalsymptoms
                           poorgrowth skinirritation vomiting other rename = (waz = waz_at_time haz=haz_at_time bmiz=bmiz_at_time
ge1_symptom=ge1_symptom_at_time abdominaldiscomfort=abdominaldiscomfort_at_time
                           anemia=anemia_at_time chronicconstipation=chronicconstipation_at_time
dentalenemaldefects=dentalenemaldefects_at_time fatigue=fatigue_at_time frequentloosestools=frequentloosestools_at_time
                           irritability=irritability_at_time neurologicalsymptoms=neurologicalsymptoms_at_time
poorgrowth=poorgrowth_at_time skinirritation=skinirritation_at_time vomiting=vomiting_at_time other=other_at_time))
             table3_part1 (in=val2 keep=maskid waz haz bmiz ge1_symptom abdominaldiscomfort anemia chronicconstipation
dentalenemaldefects fatigue frequentloosestools irritability neurologicalsymptoms
```

```
                          poorgrowth skinirritation vomiting other rename = (waz = waz_12_mo haz=haz_12_mo bmiz=bmiz_12_mo
ge1_symptom=ge1_symptom_12_mo abdominaldiscomfort=abdominaldiscomfort_12_mo
                          anemia=anemia_12_mo chronicconstipation=chronicconstipation_12_mo
dentalenemaldefects=dentalenemaldefects_12_mo fatigue=fatigue_12_mo frequentloosestools=frequentloosestools_12_mo
                          irritability=irritability_12_mo neurologicalsymptoms=neurologicalsymptoms_12_mo poorgrowth=poorgrowth_12_mo
skinirritation=skinirritation_12_mo vomiting=vomiting_12_mo other=other_12_mo))
             table3_part2 (in=val3 keep=maskid waz haz bmiz ge1_symptom abdominaldiscomfort anemia chronicconstipation
dentalenemaldefects fatigue frequentloosestools irritability neurologicalsymptoms
                          poorgrowth skinirritation vomiting other rename = (waz = waz_cda haz=haz_cda bmiz=bmiz_cda
ge1_symptom=ge1_symptom_cda abdominaldiscomfort=abdominaldiscomfort_cda
                          anemia=anemia_cda chronicconstipation=chronicconstipation_cda dentalenemaldefects=dentalenemaldefects_cda
fatigue=fatigue_cda frequentloosestools=frequentloosestools_cda
                          irritability=irritability_cda neurologicalsymptoms=neurologicalsymptoms_cda poorgrowth=poorgrowth_cda
skinirritation=skinirritation_cda vomiting=vomiting_cda other=other_cda));
       by maskid;
       if val1 and val2 and val3 then output;

proc freq data = table1;
       tables country /list;
       format country countryf.;
       title 'Table 2 - Country - CDA';

proc freq data = table1;
       tables country*ge1_symptom /nopercent nocol;
       format country countryf.;
       title 'Table 2 - Country - CDA, >= 1 Symptoms';

proc logistic data = table1;
       class country;
       model ge1_symptom = country ;
       oddsratio country;

proc freq data = table1;
       tables country;
       format country countryf.;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 2 - Country - CD';

proc freq data = table1;
       tables country*ge1_symptom /nopercent nocol;
       where celiac_disease = 1 and marsh ne .;
       format country countryf.;
       title 'Table 2 - Country - CD, >= 1 Symptoms';

proc logistic data = table1;
       class country;
       model ge1_symptom = country;
       oddsratio country;
       where celiac_disease = 1 and marsh ne .;

proc freq data = table1;
```

```
        tables sex /list;
        title 'Table 2 - Gender - CDA';

proc freq data = table1;
        tables sex*ge1_symptom /nopercent nocol;
        title 'Table 2 - Gender - CDA, >= 1 Symptoms';

proc logistic data = table1;
        class sex /ref = first;
        model ge1_symptom = sex;
        oddsratio sex;

proc freq data = table1;
        tables sex;
        where celiac_disease = 1 and marsh ne .;
        title 'Table 2 - Gender - CD';

proc freq data = table1;
        tables sex*ge1_symptom /nopercent nocol;
        where celiac_disease = 1 and marsh ne .;
        title 'Table 2 - Gender - CD, >= 1 Symptoms';

proc logistic data = table1;
        class sex /ref = first;
        model ge1_symptom = sex;
        oddsratio sex;
        where celiac_disease = 1 and marsh ne .;

proc freq data = table1;
        tables hlarg /list;
        format hlarg hlaf.;
        title 'Table 2 - HLA - CDA';

proc freq data = table1;
        tables hlarg*ge1_symptom /nopercent nocol;
        format hlarg hlaf.;
        title 'Table 2 - HLA - CDA, >= 1 Symptoms';

proc logistic data = table1;
        class hlarg;
        model ge1_symptom = hlarg ;
        oddsratio hlarg;
        format hlarg hlaf.;

proc freq data = table1 order=formatted;
        tables hlarg;
        format hlarg hlaf.;
        where celiac_disease = 1 and marsh ne .;
        title 'Table 2 - HLA - CD';

proc freq data = table1 order=formatted;
```

```
        tables hlarg*ge1_symptom /nopercent nocol;
        where celiac_disease = 1 and marsh ne .;
        format hlarg hlaf.;
        title 'Table 2 - HLA - CD, >= 1 Symptoms';

proc logistic data = table1;
        class hlarg;
        model ge1_symptom = hlarg ;
        oddsratio hlarg;
        format hlarg hlaf.;
        where celiac_disease = 1 and marsh ne .;

proc freq data = table1;
        tables celiac_fdr /list;
        title 'Table 2 - FDR with CD - CDA';

proc freq data = table1;
        tables celiac_fdr*ge1_symptom /nopercent nocol;
        title 'Table 2 - FDR with CD - HLA, >= 1 Symptoms';

proc logistic data = table1;
        class celiac_fdr;
        model ge1_symptom = celiac_fdr ;
        oddsratio celiac_fdr;

proc freq data = table1 order=formatted;
        tables celiac_fdr;
        where celiac_disease = 1 and marsh ne .;
        title 'Table 2 - FDR with CD - CD';

proc freq data = table1 order=formatted;
        tables celiac_fdr*ge1_symptom /nopercent nocol;
        where celiac_disease = 1 and marsh ne .;
        title 'Table 2 - FDR with CD - CD, >= 1 Symptoms';

proc logistic data = table1;
        class celiac_fdr;
        model ge1_symptom = celiac_fdr ;
        oddsratio celiac_fdr;
        where celiac_disease = 1 and marsh ne .;

proc means data = all_three_visits mean std;
        var waz_12_mo waz_at_time waz_cda;
        title 'Table 3 - Wt, z-score';

proc means data = all_three_visits mean std;
        var haz_12_mo haz_at_time haz_cda;
        title 'Table 3 - Height, z-score';

proc means data = all_three_visits mean std;
        var bmiz_12_mo bmiz_at_time bmiz_cda;
```

```
        title 'Table 3 - BMI, z-score';

proc freq data = all_three_visits;
        tables ge1_symptom_12_mo;
        title 'Table 3 - >= 1 symptom at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables ge1_symptom_at_time;
        title 'Table 3 - >= 1 symptom at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables ge1_symptom_cda;
        title 'Table 3 - >= 1 symptom at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables abdominaldiscomfort_12_mo;
        title 'Table 3 - Abdominal discomfort at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables abdominaldiscomfort_at_time;
        title 'Table 3 - Abdominal discomfort at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables abdominaldiscomfort_cda;
        title 'Table 3 - Abdominal discomfort at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables anemia_12_mo;
        title 'Table 3 - Anemia at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables anemia_at_time;
        title 'Table 3 - Anemia at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables anemia_cda;
        title 'Table 3 - Anemia at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables chronicconstipation_12_mo;
        title 'Table 3 - Chronic constipation at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables chronicconstipation_at_time;
        title 'Table 3 - Chronic constipation at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables chronicconstipation_cda;
        title 'Table 3 - Chronic constipation at time of persistent tTGA positivity';

proc freq data = all_three_visits;
```

```
        tables dentalenemaldefects_12_mo;
        title 'Table 3 - Dental enamel defects at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables dentalenemaldefects_at_time;
        title 'Table 3 - Dental enamel defects at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables dentalenemaldefects_cda;
        title 'Table 3 - Dental enamel defects at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables fatigue_12_mo;
        title 'Table 3 - Fatigue at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables fatigue_at_time;
        title 'Table 3 - Fatigue at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables fatigue_cda;
        title 'Table 3 - Fatigue at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables frequentloosestools_12_mo;
        title 'Table 3 - Frequent loose stools at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables frequentloosestools_at_time;
        title 'Table 3 - Frequent losoe stools at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables frequentloosestools_cda;
        title 'Table 3 - Frequent loose stools at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables irritability_12_mo;
        title 'Table 3 - Irritability at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
        tables irritability_at_time;
        title 'Table 3 - Irritability at time of initial tTGA positivity';

proc freq data = all_three_visits;
        tables irritability_cda;
        title 'Table 3 - Irritability at time of persistent tTGA positivity';

proc freq data = all_three_visits;
        tables neurologicalsymptoms_12_mo;
        title 'Table 3 - Neurological symptoms at 12 mo before initial tTGA positivity';
```

```
proc freq data = all_three_visits;
       tables neurologicalsymptoms_at_time;
       title 'Table 3 - Neurological symptoms at time of initial tTGA positivity';

proc freq data = all_three_visits;
       tables neurologicalsymptoms_cda;
       title 'Table 3 - Neurological symptoms at time of persistent tTGA positivity';

proc freq data = all_three_visits;
       tables poorgrowth_12_mo;
       title 'Table 3 - Poor growth at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
       tables poorgrowth_at_time;
       title 'Table 3 - Poor growth at time of initial tTGA positivity';

proc freq data = all_three_visits;
       tables poorgrowth_cda;
       title 'Table 3 - Poor growth at time of persistent tTGA positivity';

proc freq data = all_three_visits;
       tables skinirritation_12_mo;
       title 'Table 3 - Skin irritation at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
       tables skinirritation_at_time;
       title 'Table 3 - Skin irritation at time of initial tTGA positivity';

proc freq data = all_three_visits;
       tables skinirritation_cda;
       title 'Table 3 - Skin irritation at time of persistent tTGA positivity';

proc freq data = all_three_visits;
       tables vomiting_12_mo;
       title 'Table 3 - Vomiting at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
       tables vomiting_at_time;
       title 'Table 3 - Vomiting at time of initial tTGA positivity';

proc freq data = all_three_visits;
       tables vomiting_cda;
       title 'Table 3 - Vomiting at time of persistent tTGA positivity';

proc freq data = all_three_visits;
       tables other_12_mo;
       title 'Table 3 - Other at 12 mo before initial tTGA positivity';

proc freq data = all_three_visits;
       tables other_at_time;
       title 'Table 3 - Other at time of initial tTGA positivity';
```

```
proc freq data = all_three_visits;
       tables other_cda;
       title 'Table 3 - Other at time of persistent tTGA positivity';

proc sort data = table1;
       by descending wt_less_10_per;

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class wt_less_10_per;
       title 'Table 4 - Wt < 10th percentile - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class wt_less_10_per;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - Wt < 10th percentile - CD';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class ht_less_10_per;
       title 'Table 4 - Ht < 10th percentile - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class ht_less_10_per;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - Ht < 10th percentile - CD';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class bmi_less_10_per;
       title 'Table 4 - BMI < 10th percentile - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class bmi_less_10_per;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - BMI < 10th percentile - CD';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class ge1_symptom;
       title 'Table 4 - >= 1 Symptom - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class ge1_symptom;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - >= 1 Symptom - CD';
```

```
proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class country;
       format country countryf.;
       title 'Table 4 - Country - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class country;
       format country countryf.;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - Country - CD';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class sex;
       title 'Table 4 - Sex - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class sex;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - Sex - CD';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class hlarg;
       format hlarg hlaf.;
       title 'Table 4 - HLA - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class hlarg;
       format hlarg hlaf.;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - HLA - CD';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class celiac_fdr;
       title 'Table 4 - FDR with CD - CDA';

proc means data = table1 n median p25 p75;
       var persist_bristol1;
       class celiac_fdr;
       where celiac_disease = 1 and marsh ne .;
       title 'Table 4 - FDR with CD - CD';

proc sort data = table1;
       by maskid;
```

```
data all_survey_results;
       merge m92_data1 (in=val1)
             m92_data2 (in=val2)
             table1    (in=val3 keep=maskid);
       by maskid;
       if val1 and val2 and (persist_tga = 0 or val3) and due_num ne . then output;

data all_survey_results_2y all_survey_results_3y all_survey_results_4y other;
       set all_survey_results;
       if due_num <= 24 then output all_survey_results_2y;
       else if due_num <= 36 then output all_survey_results_3y;
       else if due_num <= 48 then output all_survey_results_4y;
       else output other;

proc sort data = all_survey_results_2y;
       by maskid due_num;

data all_survey_results_2y;
       set all_survey_results_2y;
       by maskid;
       if last.maskid then output;

proc sort data = all_survey_results_3y;
       by maskid due_num;

data all_survey_results_3y;
       set all_survey_results_3y;
       by maskid;
       if last.maskid then output;

proc sort data = all_survey_results_4y;
       by maskid due_num;

data all_survey_results_4y;
       set all_survey_results_4y;
       by maskid;
       if last.maskid then output;

proc sort data = other;
       by maskid due_num;

data other;
       set other;
       by maskid;
       if last.maskid then output;

proc freq data = all_survey_results_2y;
       tables persist_tga;

proc freq data = all_survey_results_3y;
```

```
        tables persist_tga;

proc freq data = all_survey_results_4y;
        tables persist_tga;

proc freq data = other;
        tables persist_tga;

endsas;


proc sort data = table1 nodupkey;
        by maskid due_num;

proc freq data = table1;
        tables due_num;

data table1;
        set table1;
        by maskid;
        array years(0:31) year0-year31;
        retain years;
        if first.maskid then do i = 0 to 31;
                years[i] = 0;
        end;
        if due_num = 6 then years[0] = 1;
        if due_num = 12 then years[1] = 1;
        if due_num = 18 then years[2] = 1;
        if due_num = 24 then years[3] = 1;
        if due_num = 27 then years[4] = 1;
        if due_num = 30 then years[5] = 1;
        if due_num = 33 then years[6] = 1;
        if due_num = 36 then years[7] = 1;
        if due_num = 39 then years[8] = 1;
        if due_num = 42 then years[9] = 1;
        if due_num = 45 then years[10] = 1;
        if due_num = 48 then years[11] = 1;
        if due_num = 51 then years[12] = 1;
        if due_num = 54 then years[13] = 1;
        if due_num = 57 then years[14] = 1;
        if due_num = 60 then years[15] = 1;
        if due_num = 63 then years[16] = 1;
        if due_num = 66 then years[17] = 1;
        if due_num = 69 then years[18] = 1;
        if due_num = 72 then years[19] = 1;
        if due_num = 75 then years[20] = 1;
        if due_num = 78 then years[21] = 1;
        if due_num = 81 then years[22] = 1;
        if due_num = 84 then years[23] = 1;
        if due_num = 87 then years[24] = 1;
        if due_num = 90 then years[25] = 1;
```

```
        if due_num = 93 then years[26] = 1;
        if due_num = 96 then years[27] = 1;
        if due_num = 102 then years[28] = 1;
        if due_num = 108 then years[29] = 1;
        if due_num = 111 then years[30] = 1;
        if due_num = 114 then years[31] = 1;
        if last.maskid then output;

data table1;
        set table1;
        array years(0:31) year0-year31;
        found = 0;
        first_year = .;
        do i = 0 to 31;
                if years[i] = 1 and not found then do;
                        first_year = i;
                        found = 1;
                end;
        end;

data table1;
        set table1;
        if first_year <= 3 then group = 1;
        else if first_year <= 7 then group = 2;
        else if first_year <= 11 then group = 3;

proc sort data = table1;
        by group;

proc freq data = table1;
        tables persist_tga persist_visit2;
        by group;


endsas;


data by2y_age  by3y_age by4y_age;
        set table1;
        if due_num <= 24 then output by2y_age;
        else if due_num <= 36 then output by3y_age;
        else if due_num <= 48 then output by4y_age;

proc sort data = by2y_age nodupkey;
        by maskid;

proc freq data = by2y_age;
        tables persist_tga;

proc freq data = by3y_age;
        tables persist_tga;
```

```
proc freq data = by4y_age;
        tables persist_tga;
```