Dataset Integrity Check for The Environmental Determinants of Diabetes in the Young (TEDDY) M45 Krischer

Prepared by Adam Risch
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD 20705
February 28, 2019

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_45_jkrischer_niddk_30apr2016_1.sas7bdat" dataset.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Krischer et al [1] in *Diabetes Care* in 2017. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Characteristics of HLA-eligible TEDDY children followed to age 6 years, 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 published in Table 1. The results of the replication are similar to the published results.

6 Conclusions

The NIDDK repository is confident that the TEDDY M45 data files to be distributed are a true copy of the study data.

7 References

[1] Jeffrey P. Krischer, Kristian F. Lynch, Åke Lernmark, William A. Hagopian, Marian J. Rewers, Jin-Xiong She, Jorma Toppari, Anette-G. Ziegler, Beena Akolkar, and the TEDDY Study Group. "Genetic and Environmental Interactions Modify the Risk of Diabetes-Related Autoimmunity by 6 Years of Age: The TEDDY Study". Diabetes Care 2017 Sep;40(9):1194-1202. doi: 10.2337/dc17-0238. Epub 2017 Jun 2

Table A: Variables used to replicate Table 1: Characteristics of HLA-eligible TEDDY children followed to age 6 years

Table Variable	**dataset.variable**
IA negative	m_45_jkrischer_niddk_30apr2016_1.ia_any
Any IA	m_45_jkrischer_niddk_30apr2016_1.ia_any
IAA only	m_45_jkrischer_niddk_30apr2016_1.ia_iaa_only
GADA only	m_45_jkrischer_niddk_30apr2016_1.ia_gad_only
IA2A only	m_45_jkrischer_niddk_30apr2016_1.ia_ia2a_only
Multiple IA	m_45_jkrischer_niddk_30apr2016_1.ia_multiple
Type 1 DM	m_45_jkrischer_niddk_30apr2016_1.t1d
Sex	m_45_jkrischer_niddk_30apr2016_1.male
Family History	m_45_jkrischer_niddk_30apr2016_1.fdr_proband
HLA	m_45_jkrischer_niddk_30apr2016_1.hla_dr_5grps
Country	m_45_jkrischer_niddk_30apr2016_1.country
Started probiotics	m_45_jkrischer_niddk_30apr2016_1.firstyr_probiotic_3grps
Maternal IA	m_45_jkrischer_niddk_30apr2016_1.maternal_ab_exp
Weight at 12 months, mean z score	m_45_jkrischer_niddk_30apr2016_1.wtz_12mo
Child conditions before first clinical visit (age 3 months) – Upper resp.	m_45_jkrischer_niddk_30apr2016_1.child_upper_resp_3mo
Child conditions before first clinical visit (age 3 months) – Lower resp.	m_45_jkrischer_niddk_30apr2016_1.chile_lower_resp_3mo
Child conditions before first clinical visit (age 3 months) – Diarrhea	m_45_jkrischer_niddk_30apr2016_1.child_diarrhea_episode
Child conditions before first clinical visit (age 3 months) – Rash	m_45_jkrischer_niddk_30apr2016_1.child_rash
rs689 (INS)	m_45_jkrischer_niddk_30apr2016_1.rs689_ins
rs231775 (CTLA4)	m_45_jkrischer_niddk_30apr2016_1.rs231775_ctla4
rs2476601 (PTPN22)	m_45_jkrischer_niddk_30apr2016_1. rs2476601_pptpn22
rs2816316 (RGS1)	m_45_jkrischer_niddk_30apr2016_1.rs2816316_rgs1
rs10517086	m_45_jkrischer_niddk_30apr2016_1.rs10517086
rs4948088 (COBL)	m_45_jkrischer_niddk_30apr2016_1.rs4948088_cobl
rs2292239 (ERBB3)	m_45_jkrischer_niddk_30apr2016_1.rs2292239_erbb3
rs3184504 (SH2B3)	m_45_jkrischer_niddk_30apr2016_1.rs3184504_sh2b3
rs12708716 (CLEC16A)	m_45_jkrischer_niddk_30apr2016_1.rs12708716_clec16a

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

IA negative	Manuscript (N = 7914)		DSIC (N = 7914)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	3928	49.6	3928	49.6	0	0.0
Male	3986	50.4	3986	50.4	0	0.0
Family history						
GP	7112	89.9	7112	89.9	0	0.0
FDR mother	307	3.9	307	3.9	0	0.0
FDR father	384	4.9	384	4.9	0	0.0
FDR sibling	111	1.4	111	1.4	0	0.0
HLA						
DR 3/3	1706	21.6	1706	21.6	0	0.0
DR 3/X	22	0.3	22	0.3	0	0.0
DR 3/4	3019	38.2	3019	38.1	0	0.1
DR 4/X	1609	20.3	1609	20.3	0	0.0
DR 4/4	1558	19.7	1558	19.7	0	0.0
Country						
U.S.	3428	43.3	3428	43.3	0	0.0
Finland	1658	21.0	1658	21.0	0	0.0
Germany	525	6.6	525	6.6	0	0.0
Sweden	2303	29.1	2303	29.1	0	0.0
Started probiotics						
No or age >365 days	6290	78.5	6290	79.5	0	1.0

IA negative	Manuscript (N = 7914)		DSIC (N = 7914)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes, age 28-365 days	1077	13.6	1077	13.6	0	0.0
Yes, age <28 days	547	6.9	547	6.9	0	0.0
Maternal IA						
No	7702	97.3	7702	97.3	0	0.0
Yes	212	2.7	212	2.7	0	0.0
Weight at 12 months, mean z score (SD)	-0.13	1.02	-0.13	1.02	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	1857	23.5	1857	23.5	0	0.0
Lower resp.	970	12.3	970	12.3	0	0.0
Diarrhea	740	9.4	740	9.4	0	0.0
Rash	1804	22.8	1804	22.8	0	0.0
rs689 (INS)						
No	3721	53.7	3411	53.7	310	0.0
Yes	3206	46.3	2938	46.3	268	0.0
rs231775 (CTLA4)						
No	2176	31.4	2005	31.6	171	0.2
Yes	4751	68.6	4344	68.4	407	0.2
rs2476601 (PTPN22)						
No	5550	80.1	5083	80.1	467	0.0
Yes	1376	19.9	1266	19.9	110	0.0
rs2816316 (RGS1)						
No	4271	66.9	4250	66.9	21	0.0
Yes	2110	33.1	2099	33.1	11	0.0

IA negative	Manuscript (N = 7914)		DSIC (N = 7914)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
rs10517086						
No	3281	51.4	3261	51.4	20	0.0
Yes	3100	48.6	3088	48.6	12	0.0
rs4948088 (COBL)						
No	5801	90.9	5772	90.9	29	0.0
Yes	580	9.1	577	9.1	3	0.0
rs2292239 (ERBB3)						
No	2964	46.5	2952	46.5	12	0.0
Yes	3417	53.6	3397	53.5	20	0.1
rs3184504 (SH2B3)						
No	2022	31.7	2010	31.7	12	0.0
Yes	4359	68.3	4339	68.3	20	0.0
rs12708716 (CLEC16A)						
No	2788	43.8	2780	43.8	8	0.0
Yes	3577	56.2	3569	56.2	8	0.0

Any IA	Manuscript (N = 589)		DSIC (N = 589)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	264	44.8	264	44.8	0	0.0
Male	325	55.2	325	55.2	0	0.0
Family history						
GP	465	79.0	465	78.9	0	0.1
FDR mother	31	5.3	31	5.3	0	0.0
FDR father	65	11.0	65	11.0	0	0.0
FDR sibling	28	4.8	28	4.8	0	0.0
HLA						
DR 3/3	76	12.9	76	12.9	0	0.0
DR 3/X	0	0	0	0	0	0.0
DR 3/4	298	50.6	298	50.6	0	0.0
DR 4/X	112	19.0	112	19.0	0	0.0
DR 4/4	103	17.5	103	17.5	0	0.0
Country						
U.S.	203	34.5	203	34.5	0	0.0
Finland	144	24.5	144	24.4	0	0.1
Germany	48	8.2	48	8.1	0	0.1
Sweden	194	32.9	194	32.9	0	0.0
Started probiotics						
No or age >365 days	453	76.9	453	76.9	0	0.0
Yes, age 28-365 days	100	17.0	100	17.0	0	0.0

Any IA	Manuscript (N = 589)		DSIC (N = 589)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes, age <28 days	36	6.1	36	6.1	0	0.0
Maternal IA						
No	568	96.4	568	96.4	0	0.0
Yes	21	3.6	21	3.6	0	0.0
Weight at 12 months, mean z score (SD)	0.00	1.05	0.00	1.05	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	144	24.5	144	24.4	0	0.1
Lower resp.	66	11.2	66	11.2	0	0.0
Diarrhea	52	8.9	52	8.8	0	0.1
Rash	126	21.4	126	21.4	0	0.0
rs689 (INS)						
No	381	65.6	347	64.4	34	1.2
Yes	200	34.4	192	35.6	8	1.2
rs231775 (CTLA4)						
No	174	30.0	161	29.9	13	0.1
Yes	407	70.1	378	70.1	29	0.0
rs2476601 (PTPN22)						
No	401	69.1	370	68.6	31	0.5
Yes	179	30.9	169	31.4	10	0.5
rs2816316 (RGS1)						
No	360	66.5	359	66.6	1	0.1
Yes	181	33.5	180	33.4	1	0.1
rs10517086						
No	259	47.9	259	48.1	0	0.2

Any IA	Manuscript (N = 589)		DSIC (N = 589)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes	282	52.1	280	51.9	2	0.2
rs4948088 (COBL)						
No	508	93.9	506	93.9	2	0.0
Yes	33	6.1	33	6.1	0	0.0
rs2292239 (ERBB3)						
No	200	37.0	200	37.1	0	0.1
Yes	341	63.0	339	62.9	2	0.1
rs3184504 (SH2B3)						
No	129	23.8	129	23.9	0	0.1
Yes	412	76.2	410	76.1	2	0.1
rs12708716 (CLEC16A)						
No	263	48.6	261	48.4	2	0.2
Yes	278	51.4	278	51.6	0	0.2

IAA only	Manuscript (N=252)		DSIC (N = 252)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	108	42.9	108	42.9	0	0.0
Male	144	57.1	144	57.1	0	0.0
Family history						
GP	196	77.8	196	77.8	0	0.0
FDR mother	10	4.0	10	4.0	0	0.0
FDR father	29	11.5	29	11.5	0	0.0
FDR sibling	17	6.8	17	6.7	0	0.1
HLA						
DR 3/3	17	6.8	17	6.7	0	0.1
DR 3/X	0	0	0	0	0	0.0
DR 3/4	125	49.6	125	49.6	0	0.0
DR 4/X	66	26.2	66	26.2	0	0.0
DR 4/4	44	17.5	44	17.5	0	0.0
Country						
U.S.	73	29.0	73	29.0	0	0.0
Finland	78	31.0	78	31.0	0	0.0
Germany	20	7.9	20	7.9	0	0.0
Sweden	81	32.1	81	32.1	0	0.0
Started probiotics						
No or age >365 days	191	75.8	191	75.8	0	0.0
Yes, age 28-365 days	46	18.3	46	18.3	0	0.0

IAA only	Manuscript (N=252)		DSIC (N = 252)		Diff. (N = 0)	
Yes, age <28 days	15	6.0	15	6.0	0	0.0
Maternal IA						
No	243	96.4	243	96.4	0	0.0
Yes	9	3.6	9	3.6	0	0.0
Weight at 12 months, mean z score (SD)	-0.06	1.05	-0.06	1.05	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	68	27.0	68	27.0	0	0.0
Lower resp.	27	10.7	27	10.7	0	0.0
Diarrhea	26	10.4	26	10.3	0	0.1
Rash	55	21.8	55	21.8	0	0.0
rs689 (INS)						
No	182	73.4	169	72.8	13	0.6
Yes	66	26.6	63	27.2	3	0.6
rs231775 (CTLA4)						
No	80	32.3	75	32.3	5	0.0
Yes	168	67.7	157	67.7	11	0.0
rs2476601 (PTPN22)						
No	168	67.7	155	66.8	13	0.9
Yes	80	32.3	77	33.2	3	0.9
rs2816316 (RGS1)						
No	160	68.7	160	69.0	0	0.3
Yes	73	31.3	72	31.0	1	0.3
rs10517086						
No	111	47.6	111	47.8	0	0.2
Yes	122	52.4	121	52.2	1	0.2
rs4948088 (COBL)						
No	221	94.9	220	94.8	1	0.1

IAA only	Manuscript (N=252)		DSIC (N = 252)		Diff. (N = 0)	
Yes	12	5.1	12	5.2	0	0.1
rs2292239 (ERBB3)						
No	87	37.3	87	37.5	0	0.2
Yes	146	62.7	145	62.5	1	0.2
rs3184504 (SH2B3)						
No	62	26.6	62	26.7	0	0.1
Yes	171	73.4	170	73.3	1	0.1
rs12708716 (CLEC16A)						
No	105	45.1	104	44.8	1	0.3
Yes	128	54.9	128	55.2	0	0.3

GADA only	Manuscript (N=226)		DSIC (N = 226)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	105	46.5	105	46.5	0	0.0
Male	121	53.5	121	53.5	0	0.0
Family history						
GP	183	81.0	183	81.0	0	0.0
FDR mother	12	5.3	12	5.3	0	0.0
FDR father	25	11.1	25	11.1	0	0.0
FDR sibling	6	2.7	6	2.7	0	0.0
HLA						
DR 3/3	54	23.9	54	23.9	0	0.0
DR 3/X	0	0	0	0	0	0.0
DR 3/4	111	49.1	111	49.1	0	0.0
DR 4/X	27	12.0	27	11.9	0	0.1
DR 4/4	34	15.0	34	15.0	0	0.0
Country						
U.S.	98	43.4	98	43.4	0	0.0
Finland	40	17.7	40	17.7	0	0.0
Germany	10	4.4	10	4.4	0	0.0
Sweden	78	34.5	78	34.5	0	0.0
Started probiotics						
No or age >365 days	182	80.5	182	80.5	0	0.0
Yes, age 28-365 days	33	14.6	33	14.6	0	0.0

GADA only	Manuscript (N=226)		DSIC (N = 226)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes, age <28 days	11	4.9	11	4.9	0	0.0
Maternal IA						
No	220	97.4	220	97.3	0	0.1
Yes	6	2.7	6	2.7	0	0.0
Weight at 12 months, mean z score (SD)	0.11	1.09	0.11	1.09	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	47	20.8	47	20.8	0	0.0
Lower resp.	31	13.9	31	13.7	0	0.2
Diarrhea	16	7.1	16	7.1	0	0.0
Rash	42	18.7	42	18.6	0	0.1
rs689 (INS)						
No	120	53.6	107	51.7	13	1.9
Yes	104	46.4	100	48.3	4	1.9
rs231775 (CTLA4)						
No	56	25	54	26.1	2	1.1
Yes	168	75.0	153	73.9	15	1.1
rs2476601 (PTPN22)						
No	159	71.3	147	71.0	12	0.3
Yes	64	28.7	60	29.0	4	0.3
rs2816316 (RGS1)						
No	140	67.3	139	67.1	1	0.2
Yes	68	32.7	68	32.9	0	0.2
rs10517086						
No	106	51	106	51.2	0	0.2

GADA only	Manuscript (N=226)		DSIC (N = 226)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes	102	49.0	101	48.8	1	0.2
rs4948088 (COBL)						
No	192	92.3	191	92.3	1	0.0
Yes	16	7.7	16	7.7	0	0.0
rs2292239 (ERBB3)						
No	71	34.1	71	34.3	0	0.2
Yes	137	65.9	136	65.7	1	0.2
rs3184504 (SH2B3)						
No	41	19.7	41	19.8	0	0.1
Yes	167	80.3	166	80.2	1	0.1
rs12708716 (CLEC16A)						
No	104	50.0	103	49.8	1	0.2
Yes	104	50.0	104	50.2	0	0.2

IA2A only	Manuscript (N=9)		DSIC (N = 9)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	2	22.2	2	22.2	0	0.0
Male	7	77.8	7	77.8	0	0.0
Family history						
GP	7	77.8	7	77.8	0	0.0
FDR mother	1	11.1	1	11.1	0	0.0
FDR father	0	0	0	0	0	0.0
FDR sibling	1	11.1	1	11.1	0	0.0
HLA						
DR 3/3	1	11.1	1	11.1	0	0.0
DR 3/X	0	0	0	0	0	0.0
DR 3/4	3	33.3	3	33.3	0	0.0
DR 4/X	2	22.2	2	22.2	0	0.0
DR 4/4	3	33.3	3	33.3	0	0.0
Country						
U.S.	2	22.2	2	22.2	0	0.0
Finland	2	22.2	2	22.2	0	0.0
Germany	1	11.1	1	11.1	0	0.0
Sweden	4	44.4	4	44.4	0	0.0
Started probiotics						
No or age >365 days	7	77.8	7	77.8	0	0.0
Yes, age 28-365 days	1	1.1	1	11.1	0	10.0

IA2A only	Manuscript (N=9)		DSIC (N = 9)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes, age <28 days	1	11.1	1	11.1	0	0.0
Maternal IA						
No	8	88.9	8	88.9	0	0.0
Yes	1	11.1	1	11.1	0	0.0
Weight at 12 months, mean z score (SD)	0.03	1.08	0.03	1.08	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	3	33.3	3	33.3	0	0.0
Lower resp.	2	22.2	2	22.2	0	0.0
Diarrhea	0	0	0	0	0	0.0
Rash	2	22.2	2	22.2	0	0.0
rs689 (INS)						
No	8	88.9	7	87.5	1	1.4
Yes	1	11.1	1	12.5	0	1.4
rs231775 (CTLA4)						
No	3	33.3	2	25.0	1	8.3
Yes	6	66.7	6	75.0	0	8.3
rs2476601 (PTPN22)						
No	5	55.6	4	50.0	1	5.6
Yes	4	44.4	4	50.0	0	5.6
rs2816316 (RGS1)						
No	5	62.5	5	62.5	0	0.0
Yes	3	37.5	3	37.5	0	0.0
rs10517086						
No	3	37.5	3	37.5	0	0.0

IA2A only	Manuscript (N=9)		DSIC (N = 9)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes	5	62.5	5	62.5	0	0.0
rs4948088 (COBL)						
No	8	100.0	8	100.0	0	0.0
Yes	0	0	0	0	0	0.0
rs2292239 (ERBB3)						
No	5	62.5	5	62.5	0	0.0
Yes	3	37.5	3	37.5	0	0.0
rs3184504 (SH2B3)						
No	1	12.5	1	12.5	0	0.0
Yes	7	87.5	7	87.5	0	0.0
rs12708716 (CLEC16A)						
No	4	50.0	4	50.0	0	0.0
Yes	4	50.0	4	50.0	0	0.0

Multiple IA	Manuscript (N=102)		DSIC (N = 102)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	49	48.0	49	48.0	0	0.0
Male	53	52.0	53	52.0	0	0.0
Family history						
GP	79	77.5	79	77.5	0	0.0
FDR mother	8	7.8	8	7.8	0	0.0
FDR father	11	10.8	11	10.8	0	0.0
FDR sibling	4	3.9	4	3.9	0	0.0
HLA						
DR 3/3	4	3.9	4	3.9	0	0.0
DR 3/X	0	0	0	0	0	0.0
DR 3/4	59	57.8	59	57.8	0	0.0
DR 4/X	17	16.7	17	16.7	0	0.0
DR 4/4	22	21.6	22	21.6	0	0.0
Country						
U.S.	30	29.4	30	29.4	0	0.0
Finland	24	23.5	24	23.5	0	0.0
Germany	17	16.7	17	16.7	0	0.0
Sweden	31	30.4	31	30.4	0	0.0
Started probiotics						
No or age >365 days	73	71.6	73	71.6	0	0.0
Yes, age 28-365 days	20	19.6	20	19.6	0	0.0

Multiple IA	Manuscript (N=102)		DSIC (N = 102)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes, age <28 days	9	8.8	9	8.8	0	0.0
Maternal IA						
No	97	95.1	97	95.1	0	0.0
Yes	5	4.9	5	4.9	0	0.0
Weight at 12 months, mean z score (SD)	-0.11	0.96	-0.11	0.96	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	26	26.0	26	25.5	0	0.5
Lower resp.	6	5.9	6	5.9	0	0.0
Diarrhea	10	9.8	10	9.8	0	0.0
Rash	27	26.5	27	26.5	0	0.0
rs689 (INS)						
No	71	71	64	69.6	7	1.4
Yes	29	29.0	28	30.4	1	1.4
rs231775 (CTLA4)						
No	35	35.0	30	32.6	5	2.4
Yes	65	65.0	62	67.4	3	2.4
rs2476601 (PTPN22)						
No	69	69	64	69.6	5	0.6
Yes	31	31.0	28	30.4	3	0.6
rs2816316 (RGS1)						
No	55	59.8	55	59.8	0	0.0
Yes	37	40.2	37	40.2	0	0.0
rs10517086						
No	39	42.4	39	42.4	0	0.0

Multiple IA	Manuscript (N=102)		DSIC (N = 102)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes	53	57.6	53	57.6	0	0.0
rs4948088 (COBL)						
No	87	94.6	87	94.6	0	0.0
Yes	5	5.4	5	5.4	0	0.0
rs2292239 (ERBB3)						
No	37	40.2	37	40.2	0	0.0
Yes	55	59.8	55	59.8	0	0.0
rs3184504 (SH2B3)						
No	25	27.2	25	27.2	0	0.0
Yes	67	72.8	67	72.8	0	0.0
rs12708716 (CLEC16A)						
No	50	54.4	50	54.3	0	0.1
Yes	42	45.7	42	45.7	0	0.0

Type 1 DM	Manuscript (N=172)		DSIC (N =172)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Sex						
Female	85	49.4	85	49.4	0	0.0
Male	87	50.6	87	50.6	0	0.0
Family history						
GP	118	68.6	118	68.6	0	0.0
FDR mother	12	7.0	12	7.0	0	0.0
FDR father	31	18.0	31	18.0	0	0.0
FDR sibling	11	6.4	11	6.4	0	0.0
HLA						
DR 3/3	18	10.5	18	10.5	0	0.0
DR 3/X	1	0.6	1	0.6	0	0.0
DR 3/4	97	56.4	97	56.4	0	0.0
DR 4/X	34	19.8	34	19.8	0	0.0
DR 4/4	22	12.8	22	12.8	0	0.0
Country						
U.S.	49	28.5	49	28.5	0	0.0
Finland	54	31.4	54	31.4	0	0.0
Germany	25	14.5	25	14.5	0	0.0
Sweden	44	25.6	44	25.6	0	0.0
Started probiotics						
No or age >365 days	124	72.1	124	72.1	0	0.0
Yes, age 28-365 days	33	19.2	33	19.2	0	0.0

Type 1 DM	Manuscript (N=172)		DSIC (N =172)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes, age <28 days	15	8.7	15	8.7	0	0.0
Maternal IA						
No	162	94.2	162	94.2	0	0.0
Yes	10	5.8	10	5.8	0	0.0
Weight at 12 months, mean z score (SD)	-0.06	1.05	-0.06	1.05	0	0.0
Child conditions before first clinical visit (age 3 months)						
Upper resp.	50	29.2	50	29.1	0	0.1
Lower resp.	16	9.4	16	9.3	0	0.1
Diarrhea	14	8.1	14	8.1	0	0.0
Rash	27	15.7	27	15.7	0	0.0
rs689 (INS)						
No	116	72.1	108	71.1	8	1.0
Yes	45	28.0	44	28.9	1	0.9
rs231775 (CTLA4)						
No	46	28.6	44	28.9	2	0.3
Yes	115	71.4	108	71.1	7	0.3
rs2476601 (PTPN22)						
No	103	63.9	97	63.8	6	0.1
Yes	58	26.0	55	36.2	3	10.2
rs2816316 (RGS1)						
No	107	69.9	106	69.7	1	0.2
Yes	46	30.3	46	30.3	0	0.0
rs10517086						
No	58	38.2	58	38.2	0	0.0

Type 1 DM	Manuscript (N=172)		DSIC (N =172)		Diff. (N = 0)	
	N	Pct	N	Pct	N	Pct
Yes	95	62.1	94	61.8	1	0.3
rs4948088 (COBL)						
No	144	94.1	143	94.1	1	0.0
Yes	9	5.9	9	5.9	0	0.0
rs2292239 (ERBB3)						
No	50	32.7	50	32.9	0	0.2
Yes	103	67.3	102	67.1	1	0.2
rs3184504 (SH2B3)						
No	40	26.1	40	26.3	0	0.2
Yes	113	73.9	112	73.7	1	0.2
rs12708716 (CLEC16A)						
No	74	48.4	73	48.0	1	0.4
Yes	79	51.6	79	52.0	0	0.4

Attachment A: SAS Code

```
/* Program Documentation
//* Program Name: /prj/niddk/ims_analysis/TEDDY/prog_initial_analysis/dsic_m_45.sas
//* Author: Risch
//* Date: 2/12/2019
//* Brief Description:
Performs a data set integrity check for the "M_45_Jkrischer_NIDDK_Submission".
//* Additional Description:
//* End Description:
    */

* Input files:;
libname in '/prj/niddk/ims_analysis/TEDDY/private_orig_data/M_45_Jkrischer_NIDDK_Submission/';
* End Input;

* Output files:;

* End Output;

title 'DSIC for M_45_Jkrischer_NIDDK_Submission (/prj/niddk/ims_analysis/TEDDY/prog_initial_analysis/dsic_m_45.sas)';

options nocenter validvarname=upcase missing=' ' nofmterr linesize=256 /*mprint*/;

proc format;
   value noyes
     0 = 'No'
     1 = 'Yes';

   value ia_any
     0 = 'IA negative'
     1 = 'Any IA';

   value sex
     0 = 'Female'
     1 = 'Male';

   value famhist
     0 = 'GP'
     1 = 'FDR mother'
     2 = 'FDR father'
     3 = 'FDR sibling';

   value hla
     1 = 'DR 3/3'
     2 = 'DR 3/X'
     3 = 'DR 3/4'
     4 = 'DR 4/X'
     5 = 'DR 4/4';
```

```
   value country
     1 = 'U.S.'
     2 = 'Finland'
     3 = 'Germany'
     4 = 'Sweden';

   value strtprob
     0 = 'No or age >365 days'
     1 = 'Yes, age 28-365 days'
     2 = 'Yes, age <28 days';


*** Apply labels from data dictionary to SAS datasets ***;
data in.m_45_jkrischer_niddk_30apr2016_1;
   set in.m_45_jkrischer_niddk_30apr2016_1;

   label country                = "Subject's country - 1=US, 2=FIN, 3=GER, 4=SWE"
         male                   = "Child's gender was male - 0 = no, 1 = yes"
         fdr_proband            = "First degree relative status - 0=General Pop, 1=FDR mother, 2=FDR father but not mother, 3= FDR
sibling but not parent"
         HLA_DR_5grps           = "HLA genotypes groups (1 = HLA-DR-DQ 3-2/3-2, 2 = HLA-DR-DQ 3-2/X, 3 = DR-DQ 3-2/4-8, 4= DR-DQ 4-
8/X, 5= DR-DQ 4-8/4-8)"
         HLA_DR3_grps           = "Genotype risk groups according to number of DR haplotypes (1 = HLA-DR-DQ 4-8/4-8 or X, 2 = HLA-DR-
DQ 3-2/4-8, 3= DR-DQ 3-2/3-2 or X"
         probiotics28days       = "Child was introduced to probiotics before 28 days of age - 0 = no, 1 = yes"
         firstyr_probiotic_3grps = "Child was introduced to probiotics during first year of life - 0 = no or >365 days, 1 = yes,
between 28 and 365 days inclusive, 2 = yes, less than 28 days"
         maternal_ab_exp        = "Child likely exposed to maternal islet autoantibodies - 0 = no, 1 = yes"
         wtz_12mo               = "Child's standardized weight at 12 month visit (Z-scores with weight compared to CDC growth
charts)"
         rs689_INS              = "Child had SNP rs689 in INS-23 Hph1 - 0 = no, only major alleles; 1 = yes, minor allele"
         rs231775_CTLA4         = "Child had SNP rs231775 in CTLA4 - 0 = no, only major alleles; 1 = yes, minor allele"
         rs2476601_PTPN22       = "Child had SNP rs2476601 in PTPN22 - 0 = no, only major alleles;  1 = yes, minor allele"
         rs2816316_RGS1         = "Child had SNP rs2816316 in RGS1 - 0 = no, only major alleles;  1 = yes, minor allele"
         rs10517086             = "Child had SNP rs10517086 - 0 = no, only major alleles; 1 = yes, minor allele"
         rs2292239_ERBB3        = "Child had SNP  rs2292239 in ERBB3 - 0 = no, only major alleles; 1 = yes, minor allele"
         rs3184504_SH2B3        = "Child had SNP  rs3184504 in SH2B3 - 0 = no, only major alleles; 1 = yes, minor allele"
         rs4948088_COBL         = "Child had SNP rs4948088 in COBL - 0 = no, only major alleles; 1 = yes, minor allele"
         rs12708716_CLEC16A     = "Child had SNP rs12708716_CLEC16A - 0 = no, only major alleles; 1 = yes, minor allele"
         rs3757247_BACH2        = "Child had SNP rs3757247_BACH2  - 0 = no, only major alleles; 1 = yes, minor allele"
         child_upper_resp_3mo   = "Child had an upper respiratory illness before enrollment - 0 = no, 1 yes"
         child_lower_resp_3mo   = "Child had a lower respiratory illness before enrollment - 0 = no, 1 yes"
         child_diarrhea_episode = "Child had a diarrhea episode before enrollment - 0 = no, 1 yes"
         child_rash             = "Child had a rash condition before enrollment - 0 = no, 1 yes"
         IA_any                 = "Indicates whether or not the subject had seroconverted for any of the three Islet Autoantibodies
(GADA, IA2A and IAA) by 6 years of age - 0=no, 1=yes"
         IA_gad_only            = "Indicates whether or not the subject had seroconverted for Islet Autoantibodies and had only GADA
at seroconversion - 0=no, 1=yes"
         IA_IAA_only            = "Indicates whether or not the subject had seroconverted for Islet Autoantibodies and had only IAA
at seroconversion - 0=no, 1=yes"
```

```
        IA_IA2A_only            = "Indicates whether or not the subject had seroconverted for Islet Autoantibodies and had only IA2A
at seroconversion - 0=no, 1=yes"
        IA_gad_IAA              = "Indicates whether or not the subject had seroconverted for Islet Autoantibodies and had GADA and
IAA at seroconversion - 0=no, 1=yes"
        IA_multiple             = "Indicates whether or not the subject had seroconverted for Islet Autoantibodies and had at least
two Islet autoantibodies at seroconversion - 0=no, 1=yes"
        IA_months_atrisk        = "Age (months) of child after which child was no longer observed at risk for IA. This is the age of
IA seroconversion or age when last blood draw was taken on GADA, IA2A and IAA or age 6 year visit"
        T1D                     = "0= type 1 diagnosis not recorded, 1 = diagnosed with type 1 diabetes by 6 years of age"
        T1D_months_atrisk       = "Age (months) after which the child was no longer observed at risk for type 1 diabetes. This is age
of diagnosis or age of last clinical visit for child who has no T1D or 6 year visit"
        MaskID                  = "Unique identifier for subject"
        ;

data in.m_45_jkrischer_niddk_30apr2016_2;
   set in.m_45_jkrischer_niddk_30apr2016_2;

   label visit_age_mos     = "Child's visit age (months)"
        wtz_12mo_T1D_long = "Child's standardized weight as time varying up to 12 months of age and fixed from 12 months onwards (Z-
scores with weight compared to CDC growth charts) examine on T1D"
        T1D_open_mos      = "Child's age (months) at last visit or scheduled visit if last visit was missed"
        T1D_close_mos     = "Child's age (months) at date of diagnosis for type 1 diabetes or current visit if not diagnosed or
scheduled visit if visit was missed"
        person_years_T1D  = "Person years between current visit or t1d diagnosis and last visit"
        T1D_long          = "0 = Child was not diagnosed with type 1 diabetes since last visit, 1 = Child diagnosed with type 1
diabetes, diagnosis date = T1D_close_mos"
        wtz_12mo_IA_long  = "Child's standardized weight as time varying up to 12 months of age and fixed from 12 months onwards (Z-
scores with weight compared to CDC growth charts) examine on IA"
        IA_open_mos       = "Child's age (months) at last visit for blood draw or scheduled visit if last visit for blood draw was
missed"
        IA_close_mos      = "Child's age (months) at current visit for blood draw or scheduled visit if current visit for blood draw
was missed"
        person_years_IA   = "Person years between current visit for blood draw or IA seroconversion and the last visit for blood
draw"
        IA_any_long       = "0 = Child did not seroconvert for islet autoantibodies since last visit, 1 = Child seroconverted for
islet autoantibodies since last visit, seroconversion date = IA_close_mos"
        IA_gad_only_long  = "0 = Child did not seroconvert for only GADA since last visit, 1 = Child seroconverted for only GADA
since last visit"
        IA_IAA_only_long  = "0 = Child did not seroconvert for only IAA since last visit, 1 = Child seroconverted for only IAA since
last visit"
        IA_multiple_long  = "0 = Child did not seroconvert for multiple IA since last visit, 1 = Child seroconverted for multiple IA
since last visit"
        MaskID            = "Unique identifier for subject"
        ;

data m_45_1;
   set in.m_45_jkrischer_niddk_30apr2016_1;

data m_45_2;
   set in.m_45_jkrischer_niddk_30apr2016_2;
```

```
proc contents data=_all_;

proc means data=m_45_1 n min max;
   title2 'm_45_1';

proc means data=m_45_2 n min max;
   title2 'm_45_2';

proc sort data=m_45_1 nodupkey;
   by maskid;

proc sort data=m_45_2 nodupkey;
   by maskid visit_age_mos;

ods excel file='/prj/niddk/ims_analysis/TEDDY/private_created_data/M_45_Jkrischer_table1_dsic.xlsx';

proc tabulate data=m_45_1;
   title2 "Table 1–Characteristics of HLA-eligible TEDDY children followed to age 6 years - part 1";
   class male fdr_proband hla_dr_5grps country firstyr_probiotic_3grps maternal_ab_exp /*wtz_12mo*/ child_upper_resp_3mo
child_lower_resp_3mo child_diarrhea_episode child_rash
         ia_any ia_iaa_only ia_gad_only ia_ia2a_only ia_multiple t1d / missing ;
   table all male='Sex' fdr_proband='Family history' hla_dr_5grps='HLA' country='Country' firstyr_probiotic_3grps='Started probiotics'
maternal_ab_exp='Maternal IA' /*wtz_12mo*/
         child_upper_resp_3mo='Upper resp.' child_lower_resp_3mo='Lower resp.' child_diarrhea_episode='Diarrhea' child_rash='Rash',
         (ia_any='IA negative/Any IA' ia_iaa_only='IAA only' ia_gad_only='GADA only' ia_ia2a_only='IA2A only' ia_multiple='Multiple
IA' t1d='Type 1 DM')*(n*f=5. colpctn='Pct'*f=5.1) /misstext='0' rts = 30;
   format male sex. fdr_proband famhist. hla_dr_5grps hla. country country. firstyr_probiotic_3grps strtprob. maternal_ab_exp
child_upper_resp_3mo child_lower_resp_3mo child_diarrhea_episode child_rash noyes.
          ia_any ia_any. ia_iaa_only ia_gad_only ia_ia2a_only ia_multiple t1d noyes.;
run;

proc tabulate data=m_45_1;
   title2 "Table 1–Characteristics of HLA-eligible TEDDY children followed to age 6 years - part 2";
   class ia_any ia_iaa_only ia_gad_only ia_ia2a_only ia_multiple t1d;
   var wtz_12mo;
   table wtz_12mo='Weight at 12 months, mean z score (SD)', (ia_any='IA negative/Any IA' ia_iaa_only='IAA only' ia_gad_only='GADA
only' ia_ia2a_only='IA2A only' ia_multiple='Multiple IA' t1d='Type 1 DM')*(mean*f=6.2 std*f=6.2) /misstext='0' rts = 30;
   format ia_any ia_any. ia_iaa_only ia_gad_only ia_ia2a_only ia_multiple t1d noyes.;
run;

proc tabulate data=m_45_1;
   title2 "Table 1–Characteristics of HLA-eligible TEDDY children followed to age 6 years - part 3";
   class rs689_ins rs231775_ctla4 rs2476601_ptpn22 rs2816316_rgs1 rs10517086 rs4948088_cobl rs2292239_erbb3 rs3184504_sh2b3
rs12708716_clec16a
         ia_any ia_iaa_only ia_gad_only ia_ia2a_only ia_multiple t1d;
   table rs689_ins='rs689 (INS)' rs231775_ctla4='rs231775 (CTLA4)' rs2476601_ptpn22='rs2476601 (PTPN22)' rs2816316_rgs1='rs2816316
(RGS1)' rs10517086='rs10517086' rs4948088_cobl='rs4948088 (COBL)' rs2292239_erbb3='rs2292239 (ERBB3)'
         rs3184504_sh2b3='rs3184504 (SH2B3)' rs12708716_clec16a='rs12708716 (CLEC16A)',
         (ia_any='IA negative/Any IA' ia_iaa_only='IAA only' ia_gad_only='GADA only' ia_ia2a_only='IA2A only' ia_multiple='Multiple
IA' t1d='Type 1 DM')*(n*f=5. colpctn='Pct'*f=5.1) /misstext='0' rts = 30;
```

```
   format rs689_ins rs231775_ctla4 rs2476601_ptpn22 rs2816316_rgs1 rs10517086 rs4948088_cobl rs2292239_erbb3 rs3184504_sh2b3
rs12708716_clec16a ia_iaa_only ia_gad_only ia_ia2a_only ia_multiple t1d noyes. ia_any ia_any.;
run;

ods excel close;
```