

# Dataset Integrity Check for The Environmental Determinants of Diabetes in the Young (TEDDY) Pub80 Krischer

**Prepared by Allyson Mateja**

**IMS Inc.**

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

**December 13, 2016**

## Contents

1 Standard Disclaimer .....	2
2 Study Background .....	2
3 Archived Datasets .....	2
4 Statistical Methods .....	2
5 Results .....	3
6 Conclusions .....	3
7 References .....	3
Table A: Variables used to replicate Table 1: Incidence of IAA only, GADA only, and a combination of IAA and GADA and time to seroconversion.....	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	4
Attachment A: SAS Code.....	7

## **1 Standard Disclaimer**

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless NIDDK Repository staff suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

## **2 Study Background**

The TEDDY study was designed to follow children with and without a family history of 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\_80\_jkrischer\_niddk\_30apr2014.sas7bdat” dataset.

## **4 Statistical Methods**

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

## 5 Results

For Table 1 in the publication [1], Incidence of IAA only, GADA only, and a combination of IAA and GADA and time to seroconversion, 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.

## 6 Conclusions

The NIDDK repository is confident that the TEDDY M80 data files to be distributed are a true copy of the study data.

## 7 References

[1] Krischer, J.P., Lynch, K.F., Schatz, D.A., Ilonen, J., Lernmark, A., Hagopian, W.A., Rewers, M.J., She, J., Simell, O.G., Toppari, J., Ziegler, A., Akolkar, B., Bonifacio, E., and the TEDDY study group. "The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY Study". *Diabetologia* (2015) 58(5):980-987.

**Table A:** Variables used to replicate Table 1: Incidence of IAA only, GADA only, and a combination of IAA and GADA and time to seroconversion

Table Variable	Variable
IAA only at seroconversion	persist72_miaa_only
GADA only at seroconversion	persist72_gad_only
IAA and GADA at seroconversion	persist72_gad_miaa
FDR	fdr
HLA-DR genotype	hla_5grps
Country	country
Time to seroconversion	persist72_atrisk_age

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

IAA only at seroconversion (n=240)

Participant Groups	No./1,000 person-years Manuscript	No./1,000 person-years DSIC	Diff.	Time (months) to seroconversion, median (IQR) Manuscript	Time (months) to seroconversion, median (IQR) DSIC	Diff.
All (n=8,503)	7.0	7.0	0	18.4 (10.5-30.6)	18.4 (10.5-30.6)	0 (0-0)
FDR						
No (n=7,584)	6.1	6.1	0	18.4 (10.4-30.3)	18.4 (10.3-30.3)	0 (0-0)
Yes (n=919)	14.5	14.6	0.1	18.7 (11.9-34.7)	18.7 (11.9-34.7)	0 (0-0)
HLA-DR genotype						
DR3/4 (n=3,318)	9.1	9.1	0	18.2 (10.0-27.5)	18.2 (10.0-27.5)	0 (0-0)
DR4/4 (n=1,660)	6.3	6.3	0	22.9 (10.0-37.1)	22.8 (10.0-37.1)	0.1 (0-0)
DR4/8 (n=1,470)	8.8	8.8	0	16.1 (10.2-29.5)	16.1 (10.2-29.5)	0 (0-0)
DR3/3 (n=1,781)	2.0	1.9	0.1	25.7 (15.4-43.8)	25.7 (15.3-43.8)	0 (0.1-0)
FDR HLA (n=274)	11.0	11.0	0	19.1 (15.3-40.2)	19.1 (15.3-40.2)	0 (0-0)
Country						
USA (n=3,627)	5.2	5.2	0	21.4 (13.1-30.0)	21.4 (13.1-30.0)	0 (0-0)
Finland (n=1,805)	9.6	9.6	0	15.3 (9.2-27.1)	15.3 (9.2-27.1)	0 (0-0)
Sweden (n=2,497)	7.5	7.2	0.3	19.0 (10.2-35.7)	19.0 (10.3-35.7)	0 (0.1-0)
Germany (n=574)	9.0	9.0	0	16.3 (9.5-35.3)	16.3 (9.5-35.3)	0 (0-0)

GADA only at seroconversion (n=207)

Participant Groups	No./1,000 person-years Manuscript	No./1,000 person-years DSIC	Diff.	Time (months) to seroconversion, median (IQR) Manuscript	Time (months) to seroconversion, median (IQR) DSIC	Diff.
All (n=8,503)	6.1	6.1	0	33.0 (19.2-48.4)	33.0 (19.2-48.4)	0 (0-0)
FDR						
No (n=7,584)	5.5	5.5	0	33.0 (19.4-48.1)	33.0 (19.4-48.1)	0 (0-0)
Yes (n=919)	10.6	10.7	0.1	28.2 (18.5-49.8)	28.2 (18.5-49.8)	0 (0-0)
HLA-DR genotype						
DR3/4 (n=3,318)	7.6	7.6	0	35.4 (21.1-50.8)	35.4 (21.1-50.8)	0 (0-0)
DR4/4 (n=1,660)	4.6	4.6	0	26.9 (21.3-43.1)	26.9 (21.3-43.1)	0 (0-0)
DR4/8 (n=1,470)	3.6	3.6	0	33.0 (19.1-48.4)	33.0 (19.1-48.4)	0 (0-0)
DR3/3 (n=1,781)	7.0	7.0	0	27.8 (16.7-45.3)	27.8 (16.7-45.3)	0 (0-0)
FDR HLA (n=274)	3.4	3.4	0	37.6 (29.3-55.6)	37.6 (29.3-55.6)	0 (0-0)
Country						
USA (n=3,627)	6.5	6.5	0	31.3 (19.4-47.8)	31.3 (19.4-47.8)	0 (0-0)
Finland (n=1,805)	4.5	4.5	0	32.8 (22.1-48.4)	32.8 (22.1-48.4)	0 (0-0)
Sweden (n=2,497)	7.0	6.8	0.2	33.0 (19.3-48.6)	33.0 (19.3-48.6)	0 (0-0)
Germany (n=574)	5.5	5.5	0	28.2 (13.0-50.8)	28.2 (13.0-50.8)	0 (0-0)

IAA and GADA at seroconversion (n=76)

Participant Groups	No./1,000 person-years Manuscript	No./1,000 person-years DSIC	Diff.	Time (months) to seroconversion, median (IQR) Manuscript	Time (months) to seroconversion, median (IQR) DSIC	Diff.
All (n=8,503)	2.2	2.2	0	24.1 (12.9-34.8)	24.1 (12.9-34.8)	0 (0-0)
FDR						
No (n=7,584)	1.9	1.9	0	27.0 (12.9-35.8)	27.0 (12.9-35.8)	0 (0-0)
Yes (n=919)	4.6	4.7	0.1	22.4 (12.2-27.9)	22.4 (12.2-27.9)	0 (0-0)
HLA-DR genotype						
DR3/4 (n=3,318)	3.7	3.7	0	24.2 (12.9-33.2)	24.2 (12.9-33.1)	0 (0-0.1)
DR4/4 (n=1,660)	1.5	1.5	0	31.4 (19.3-39.1)	31.4 (19.3-39.1)	0 (0-0)
DR4/8 (n=1,470)	1.0	1.0	0	27.0 (20.9-33.7)	27.0 (20.9-33.7)	0 (0-0)
DR3/3 (n=1,781)	0.4	0.4	0	12.8 (12.4-18.3)	12.8 (12.4-18.3)	0 (0-0)
FDR HLA (n=274)	6.8	6.7	0.1	22.8 (11.0-33.1)	22.8 (11.0-33.1)	0 (0-0)
Country						
USA (n=3,627)	1.7	1.7	0	24.2 (15.2-33.7)	24.2 (15.2-33.7)	0 (0-0)

Participant Groups	No./1,000 person-years Manuscript	No./1,000 person-years DSIC	Diff.	Time (months) to seroconversion, median (IQR) Manuscript	Time (months) to seroconversion, median (IQR) DSIC	Diff.
Finland (n=1,805)	2.0	2.0	0	27.5 (21.9-33.3)	27.5 (21.9-33.3)	0 (0-0)
Sweden (n=2,497)	2.4	2.2	0.2	26.8 (12.3-40.5)	26.8 (12.3-40.5)	0 (0-0)
Germany (n=574)	6.5	6.5	0	22.1 (12.4-23.5)	22.1 (12.4-23.5)	0 (0-0)

## Attachment A: SAS Code

```
*** TEDDY M80 Dataset Integrity Check;
*** Programmer: Allyson Mateja;
*** Date: 11/23/16;

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

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

    value hlaf 1 = 'DR3/4'
              2 = 'DR4/4'
              3 = 'DR4/8'
              4 = 'FDR HLA'
              5 = 'DR3/3';

libname m80data '/prj/niddk/ims_analysis/TEDDY/private_orig_data/m_80_jkrischer_niddk_submission';

data m_80_data;
    set m80data.m_80_jkrischer_niddk_30apr2014;

proc contents data = m_80_data;

data table1;
    set m_80_data;
    time_yrs = persist72_atrisk_age/12;
    if exclude ne 1;

proc freq data = table1;
    tables fdr ;
    format fdr yesnof.;
    title 'Table 1 - FDR';

proc freq data = table1;
    tables hla_5grps;
    format hla_5grps hlaf.;
    title 'Table 1 - HLA-DR genotype';

proc freq data = table1;
    tables country ;
    format country countryf.;
    title 'Table 1 - Country';

proc freq data = table1;
```



```

tables fdr*persist72_miaa_only /out = freq_data_fdr_iaa noprint ;

proc freq data = table1;
  tables fdr*persist72_gad_only /out = freq_data_fdr_gad noprint ;

proc freq data = table1;
  tables fdr*persist72_gad_miaa /out = freq_data_fdr_iaa_gad noprint ;

proc freq data = table1;
  tables hla_5grps*persist72_miaa_only /out = freq_data_hla_iaa noprint ;

proc freq data = table1;
  tables hla_5grps*persist72_gad_only /out = freq_data_hla_gad noprint ;

proc freq data = table1;
  tables hla_5grps*persist72_gad_miaa /out = freq_data_hla_iaa_gad noprint ;

proc freq data = table1;
  tables country*persist72_miaa_only /out = freq_data_country_iaa noprint ;

proc freq data = table1;
  tables country*persist72_gad_only /out = freq_data_country_gad noprint ;

proc freq data = table1;
  tables country*persist72_gad_miaa /out = freq_data_country_iaa_gad noprint ;

proc means data = table1 n sum noprint;
  var time_yrs;
  output out = total sum=sum;

proc means data = table1 n sum noprint;
  var time_yrs;
  where fdr = 0;
  output out = total_no_fdr sum=sum;

proc means data = table1 n sum noprint;
  var time_yrs;
  where fdr = 1;
  output out = total_yes_fdr sum=sum;

proc means data = table1 n sum noprint;
  var time_yrs;
  where hla_5grps = 1;
  output out = total_hla1 sum=sum;

proc means data = table1 n sum noprint;
  var time_yrs;
  where hla_5grps = 2;
  output out = total_hla2 sum=sum;

proc means data = table1 n sum noprint;

```

```

        var time_yrs;
        where hla_5grps = 3;
        output out = total_hla3 sum=sum;

proc means data = table1 n sum noprint;
    var time_yrs;
    where hla_5grps = 4;
    output out = total_hla4 sum=sum;

proc means data = table1 n sum noprint;
    var time_yrs;
    where hla_5grps = 5;
    output out = total_hla5 sum=sum;

proc means data = table1 n sum noprint;
    var time_yrs;
    where country = 1;
    output out = total_us sum=sum;

proc means data = table1 n sum noprint;
    var time_yrs;
    where country = 2;
    output out = total_finland sum=sum;

proc means data = table1 n sum noprint;
    var time_yrs;
    where country = 3;
    output out = total_germany sum=sum;

proc means data = table1 n sum noprint;
    var time_yrs;
    where country = 4;
    output out = total_sweden sum=sum;

proc freq data = table1;
    tables persist72_miaa_only /out=freq_data;
    format persist72_miaa_only yesnof.;
    title 'Table 1 - IAA only at seroconversion';

data freq_data;
    set freq_data;
    if persist72_miaa_only = 1;

data person_years_total;
    merge freq_data
          total;
    person_years = count/(sum/1000);

proc print data = person_years_total;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - All';

```

```

proc freq data = table1;
    tables persist72_gad_only /out=freq_data;
    format persist72_gad_only yesnof.;
    title 'Table 1 - GADA only at seroconversion';

data freq_data;
    set freq_data;
    if persist72_gad_only = 1;

data person_years_total;
    merge freq_data
        total;
    person_years = count/(sum/1000);

proc print data = person_years_total;
    var person_years;
    title 'Table 1 - GADA only No. /1,000 person-years - All';

proc freq data = table1;
    tables persist72_gad_miaa /out=freq_data;
    format persist72_gad_miaa yesnof.;
    title 'Table 1 - IAA and GADA at seroconversion';

data freq_data;
    set freq_data;
    if persist72_gad_miaa = 1;

data person_years_total;
    merge freq_data
        total;
    person_years = count/(sum/1000);

proc print data = person_years_total;
    var person_years;
    title 'Table 1 - IAA and GADA No. /1,000 person-years - All';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    where persist72_miaa_only = 1;
    title 'Table 1 - Time to seroconversion for IAA only - All';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    class fdr;
    format fdr yesnof.;
    where persist72_miaa_only = 1;
    title 'Table 1 - Time to seroconversion for IAA only - by FDR';

data freq_data_no;
    set freq_data_fdr_iaa;

```

```

        if fdr = 0 and persist72_miaa_only = 1;

data person_years_fdr_no;
    merge freq_data_no
        total_no_fdr;
    person_years = count/(sum/1000);

proc print data = person_years_fdr_no;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - FDR=No';

data freq_data_yes;
    set freq_data_fdr_iaa;
    if fdr = 1 and persist72_miaa_only = 1;

data person_years_fdr_yes;
    merge freq_data_yes
        total_yes_fdr;
    person_years = count/(sum/1000);

proc print data = person_years_fdr_yes;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - FDR=Yes';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    class hla_5grps;
    format hla_5grps h1af.;
    where persist72_miaa_only = 1;
    title 'Table 1 - Time to seroconversion for IAA only - by HLA-DR genotype';

data freq_data_hla1;
    set freq_data_hla_iaa;
    if hla_5grps = 1 and persist72_miaa_only = 1;

data person_years_hla1;
    merge freq_data_hla1
        total_hla1;
    person_years = count/(sum/1000);

proc print data = person_years_hla1;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - HLA=DR3/4';

data freq_data_hla2;
    set freq_data_hla_iaa;
    if hla_5grps = 2 and persist72_miaa_only = 1;

data person_years_hla2;
    merge freq_data_hla2
        total_hla2;

```

```

        person_years = count/(sum/1000);

proc print data = person_years_hla2;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - HLA=DR4/4';

data freq_data_hla3;
    set freq_data_hla_iaa;
    if hla_5grps = 3 and persist72_miaa_only = 1;

data person_years_hla3;
    merge freq_data_hla3
          total_hla3;
    person_years = count/(sum/1000);

proc print data = person_years_hla3;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - HLA=DR4/8';

data freq_data_hla5;
    set freq_data_hla_iaa;
    if hla_5grps = 5 and persist72_miaa_only = 1;

data person_years_hla5;
    merge freq_data_hla5
          total_hla5;
    person_years = count/(sum/1000);

proc print data = person_years_hla5;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - HLA=DR3/3';

data freq_data_hla4;
    set freq_data_hla_iaa;
    if hla_5grps = 4 and persist72_miaa_only = 1;

data person_years_hla4;
    merge freq_data_hla4
          total_hla4;
    person_years = count/(sum/1000);

proc print data = person_years_hla4;
    var person_years;
    title 'Table 1 - IAA only No. /1,000 person-years - HLA=FDR HLA';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    class country;
    format country countryf.;
    where persist72_miaa_only = 1;
    title 'Table 1 - Time to seroconversion for IAA only - by Country';

```

```

data freq_data_us;
  set freq_data_country_iaa;
  if country = 1 and persist72_miaa_only = 1;

data person_years_us;
  merge freq_data_us
        total_us;
  person_years = count/(sum/1000);

proc print data = person_years_us;
  var person_years;
  title 'Table 1 - IAA only No. /1,000 person-years - Country=US';

data freq_data_finland;
  set freq_data_country_iaa;
  if country = 2 and persist72_miaa_only = 1;

data person_years_finland;
  merge freq_data_finland
        total_finland;
  person_years = count/(sum/1000);

proc print data = person_years_finland;
  var person_years;
  title 'Table 1 - IAA only No. /1,000 person-years - Country=Finland';

data freq_data_germany;
  set freq_data_country_iaa;
  if country = 3 and persist72_miaa_only = 1;

data person_years_germany;
  merge freq_data_germany
        total_germany;
  person_years = count/(sum/1000);

proc print data = person_years_germany;
  var person_years;
  title 'Table 1 - IAA only No. /1,000 person-years - Country=Germany';

data freq_data_sweden;
  set freq_data_country_iaa;
  if country = 4 and persist72_miaa_only = 1;

data person_years_sweden;
  merge freq_data_sweden
        total_sweden;
  person_years = count/(sum/1000);

proc print data = person_years_sweden;
  var person_years;

```

```

        title 'Table 1 - IAA only No. /1,000 person-years - Country=Sweden';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    where persist72_gad_only = 1;
    title 'Table 1 - Time to seroconversion for GADA only - All';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    class fdr;
    format fdr yesnof.;
    where persist72_gad_only = 1;
    title 'Table 1 - Time to seroconversion for GADA only - by FDR';

data freq_data_no;
    set freq_data_fdr_gad;
    if fdr = 0 and persist72_gad_only = 1;

data person_years_fdr_no;
    merge freq_data_no
          total_no_fdr;
    person_years = count/(sum/1000);

proc print data = person_years_fdr_no;
    var person_years;
    title 'Table 1 - GADA only No. /1,000 person-years - FDR=No';

data freq_data_yes;
    set freq_data_fdr_gad;
    if fdr = 1 and persist72_gad_only = 1;

data person_years_fdr_yes;
    merge freq_data_yes
          total_yes_fdr;
    person_years = count/(sum/1000);

proc print data = person_years_fdr_yes;
    var person_years;
    title 'Table 1 - GADA only No. /1,000 person-years - FDR=Yes';

proc means data = table1 n median p25 p75;
    var persist72_atrisk_age;
    class hla_5grps;
    format hla_5grps hlaf.;
    where persist72_gad_only = 1;
    title 'Table 1 - Time to seroconversion for GADA only - by HLA-DR genotype';

data freq_data_hla1;
    set freq_data_hla_gad;
    if hla_5grps = 1 and persist72_gad_only = 1;

```

```

data person_years_hla1;
  merge freq_data_hla1
        total_hla1;
  person_years = count/(sum/1000);

proc print data = person_years_hla1;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - HLA=DR3/4';

data freq_data_hla2;
  set freq_data_hla_gad;
  if hla_5grps = 2 and persist72_gad_only = 1;

data person_years_hla2;
  merge freq_data_hla2
        total_hla2;
  person_years = count/(sum/1000);

proc print data = person_years_hla2;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - HLA=DR4/4';

data freq_data_hla3;
  set freq_data_hla_gad;
  if hla_5grps = 3 and persist72_gad_only = 1;

data person_years_hla3;
  merge freq_data_hla3
        total_hla3;
  person_years = count/(sum/1000);

proc print data = person_years_hla3;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - HLA=DR4/8';

data freq_data_hla5;
  set freq_data_hla_gad;
  if hla_5grps = 5 and persist72_gad_only = 1;

data person_years_hla5;
  merge freq_data_hla5
        total_hla5;
  person_years = count/(sum/1000);

proc print data = person_years_hla5;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - HLA=DR3/3';

data freq_data_hla4;
  set freq_data_hla_gad;
  if hla_5grps = 4 and persist72_gad_only = 1;

```



```

data person_years_hla4;
  merge freq_data_hla4
        total_hla4;
  person_years = count/(sum/1000);

proc print data = person_years_hla4;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - HLA=FDR HLA';

proc means data = table1 n median p25 p75;
  var persist72_atrisk_age;
  class country;
  format country countryf.;
  where persist72_gad_only = 1;
  title 'Table 1 - Time to seroconversion for GADA only - by Country';

data freq_data_us;
  set freq_data_country_gad;
  if country = 1 and persist72_gad_only = 1;

data person_years_us;
  merge freq_data_us
        total_us;
  person_years = count/(sum/1000);

proc print data = person_years_us;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - Country=US';

data freq_data_finland;
  set freq_data_country_gad;
  if country = 2 and persist72_gad_only = 1;

data person_years_finland;
  merge freq_data_finland
        total_finland;
  person_years = count/(sum/1000);

proc print data = person_years_finland;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - Country=Finland';

data freq_data_germany;
  set freq_data_country_gad;
  if country = 3 and persist72_gad_only = 1;

data person_years_germany;
  merge freq_data_germany
        total_germany;
  person_years = count/(sum/1000);

```

```

proc print data = person_years_germany;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - Country=Germany';

data freq_data_sweden;
  set freq_data_country_gad;
  if country = 4 and persist72_gad_only = 1;

data person_years_sweden;
  merge freq_data_sweden
        total_sweden;
  person_years = count/(sum/1000);

proc print data = person_years_sweden;
  var person_years;
  title 'Table 1 - GADA only No. /1,000 person-years - Country=Sweden';

proc means data = table1 n median p25 p75;
  var persist72_atrisk_age;
  where persist72_gad_miaa = 1;
  title 'Table 1 - Time to seroconversion for IAA and GADA - All';

proc means data = table1 n median p25 p75;
  var persist72_atrisk_age;
  class fdr;
  format fdr yesnof.;
  where persist72_gad_miaa = 1;
  title 'Table 1 - Time to seroconversion for IAA and GADA - by FDR';

data freq_data_no;
  set freq_data_fdr_iaa_gad;
  if fdr = 0 and persist72_gad_miaa = 1;

data person_years_fdr_no;
  merge freq_data_no
        total_no_fdr;
  person_years = count/(sum/1000);

proc print data = person_years_fdr_no;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - FDR=No';

data freq_data_yes;
  set freq_data_fdr_iaa_gad;
  if fdr = 1 and persist72_gad_miaa = 1;

data person_years_fdr_yes;
  merge freq_data_yes
        total_yes_fdr;
  person_years = count/(sum/1000);

```

```

proc print data = person_years_fdr_yes;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - FDR=Yes';

proc means data = table1 n median p25 p75;
  var persist72_atrisk_age;
  class hla_5grps;
  format hla_5grps hla5.;
  where persist72_gad_miaa = 1;
  title 'Table 1 - Time to seroconversion for IAA and GADA - by HLA-DR genotype';

data freq_data_hla1;
  set freq_data_hla_iaa_gad;
  if hla_5grps = 1 and persist72_gad_miaa = 1;

data person_years_hla1;
  merge freq_data_hla1
        total_hla1;
  person_years = count/(sum/1000);

proc print data = person_years_hla1;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - HLA=DR3/4';

data freq_data_hla2;
  set freq_data_hla_iaa_gad;
  if hla_5grps = 2 and persist72_gad_miaa = 1;

data person_years_hla2;
  merge freq_data_hla2
        total_hla2;
  person_years = count/(sum/1000);

proc print data = person_years_hla2;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - HLA=DR4/4';

data freq_data_hla3;
  set freq_data_hla_iaa_gad;
  if hla_5grps = 3 and persist72_gad_miaa = 1;

data person_years_hla3;
  merge freq_data_hla3
        total_hla3;
  person_years = count/(sum/1000);

proc print data = person_years_hla3;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - HLA=DR4/8';

```

```

data freq_data_hla5;
  set freq_data_hla_iaa_gad;
  if hla_5grps = 5 and persist72_gad_miaa = 1;

data person_years_hla5;
  merge freq_data_hla5
        total_hla5;
  person_years = count/(sum/1000);

proc print data = person_years_hla5;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - HLA=DR3/3';

data freq_data_hla4;
  set freq_data_hla_iaa_gad;
  if hla_5grps = 4 and persist72_gad_miaa = 1;

data person_years_hla4;
  merge freq_data_hla4
        total_hla4;
  person_years = count/(sum/1000);

proc print data = person_years_hla4;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - HLA=FDR HLA';

proc means data = table1 n median p25 p75;
  var persist72_atrisk_age;
  class country;
  format country countryf.;
  where persist72_gad_miaa = 1;
  title 'Table 1 - Time to seroconversion for IAA and GADA - by Country';

data freq_data_us;
  set freq_data_country_iaa_gad;
  if country = 1 and persist72_gad_miaa = 1;

data person_years_us;
  merge freq_data_us
        total_us;
  person_years = count/(sum/1000);

proc print data = person_years_us;
  var person_years;
  title 'Table 1 - IAA and GADA No. /1,000 person-years - Country=US';

data freq_data_finland;
  set freq_data_country_iaa_gad;
  if country = 2 and persist72_gad_miaa = 1;

data person_years_finland;

```

```

merge freq_data_finland
      total_finland;
person_years = count/(sum/1000);

proc print data = person_years_finland;
var person_years;
title 'Table 1 - IAA and GADA No. /1,000 person-years - Country=Finland';

data freq_data_germany;
set freq_data_country_iaa_gad;
if country = 3 and persist72_gad_miaa = 1;

data person_years_germany;
merge freq_data_germany
      total_germany;
person_years = count/(sum/1000);

proc print data = person_years_germany;
var person_years;
title 'Table 1 - IAA and GADA No. /1,000 person-years - Country=Germany';

data freq_data_sweden;
set freq_data_country_iaa_gad;
if country = 4 and persist72_gad_miaa = 1;

data person_years_sweden;
merge freq_data_sweden
      total_sweden;
person_years = count/(sum/1000);

proc print data = person_years_sweden;
var person_years;
title 'Table 1 - IAA and GADA No. /1,000 person-years - Country=Sweden';

```