

Dataset Integrity Check for DCCT/EDIC CVD Risk Factors Data Files

Prepared by Allyson Mateja

IMS, Inc.

3901 Calverton Blvd, Suite 200, Calverton, MD 20705

February 22, 2017

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 Figure 1	4
Figure A: Comparison of values computed in integrity check to reference article Figure 1 values	4
Table B: Variables used to replicate Table 1: Baseline characteristics of DCCT/EDIC participants according to the presence or absence of any-CVD over the course of DCCT/EDIC	5
Table D: Comparison of values computed in integrity check to reference article Table 1 values	6
Attachment A: SAS Code	11

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 Epidemiology of Diabetes Interventions and Complications (EDIC) study was initiated as follow-up to examine the long-term effects of the original DCCT interventions on diabetic complications such as cardiovascular events and advanced retinal and renal disease. Over 90 percent of participants from the DCCT study were followed by the EDIC study. Similar to the DCCT study, glycosylated hemoglobin values, fasting lipid levels, serum creatinine values, and other risk factors for cardiovascular disease were measured at different intervals for participants. Cardiovascular complications were assessed with standardized means and classified by an independent committee. The EDIC study has found that intensive diabetes therapy reduced risk of cardiovascular disease in patients with type 1 diabetes and that the differences in outcomes between the intensive and conventional therapy groups persist after long-term study.

3 Archived Datasets

The SAS data file, as provided by the Data Coordinating Center (DCC), are located in the data package. For this replication, variables were taken from the SAS files, nih_cvd30yr_ev_366.sas7bdat, nih_cvd30yr_1441.sas7bdat, and cvd_riskfactors_rel.sas7bdat.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Lachin, et al [1] in Diabetes in May 2016. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Figure 1 in the publication [1], Table A lists the variables that can be used in the replication, and Figure A compares the results calculated from the archived data file to the results published in Figure 1. The results of the replication are almost an exact match to the results in the publication [1].

For Table 1 in the publication [1], Baseline characteristics of DCCT/EDIC participants according to the presence or absence of any-CVD over the course of DCCT/EDIC, Table B lists the variables that can be used in the replication, and Table C 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 results in publication [1].

6 Conclusions

The NIDDK repository is confident that the DCCT/EDIC CVD data files to be distributed are a true copy to the manuscript data.

7 References

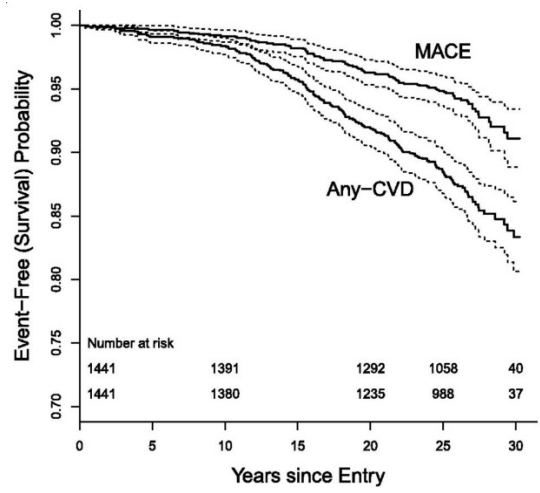
[1] Lachin JM, et al.; DCCT/EDIC Research Group. Risk Factors for Cardiovascular Disease in Type 1 Diabetes. *Diabetes* 2016;65:1370-1379.

Table A: Variables used to replicate Figure 1

Figure Variable	dataset.variable
Years since Entry (Any-CVD)	nih_cvd30yr_1441.carvtime1
Years since Entry (MACE)	nih_cvd30yr_1441.hardtime1
Any-CVD	nih_cvd30yr_1441.carv
MACE	nih_cvd30yr_1441.hard

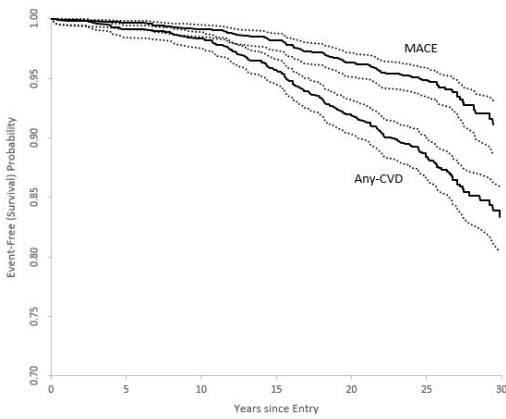
Figure A: Comparison of values computed in integrity check to reference article Figure 1 values

Manuscript:



Kaplan-Meier survival (event-free) curves (solid lines), and 95% Cis (dashed lines) for MACE (top line) and any-CVD (bottom line), with number of subjects at risk for each at 10, 20, and 30 years of follow-up. The number at risk beyond 23 years declines as a function of staggered entry into the study from 1983 to 1989.

DSIC:



	Year 0 Manuscript	Year 0 DSIC	Diff.	Year 10 Manuscript	Year 10 DSIC	Diff.	Year 20 Manuscript	Year 20 DSIC	Diff.
Number at Risk									
MACE	1441	1441	0	1391	1391	0	1292	1292	0
Any-CVD	1441	1441	0	1380	1380	0	1235	1235	0

	Year 25 Manuscript	Year 25 DSIC	Diff.	Year 30 Manuscript	Year 30 DSIC	Diff.
Number at Risk						
MACE	1058	1058	0	40	40	0
Any-CVD	988	988	0	37	37	0

Table B: Variables used to replicate Table 1: Baseline characteristics of DCCT/EDIC participants according to the presence or absence of any-CVD over the course of DCCT/EDIC

Table Variable	dataset.variable
Any-CVD	nih_cvd30yr_ev_366.cvd_event
Treatment group	cvd_riskfactors_rel.group
Cohort	cvd_riskfactors_rel.retbase
Sex	cvd_riskfactors_rel.sex
Age	cvd_riskfactors_rel.age
Adult vs. adolescent	cvd_riskfactors_rel.adult
Weight	cvd_riskfactors_rel.weight
BMI	cvd_riskfactors_rel.bmi
Smoking	cvd_riskfactors_rel.smokes
Alcohol	cvd_riskfactors_rel.drinks
Exercise	cvd_riskfactors_rel.exercise
Family history of hypertension	cvd_riskfactors_rel.famht
Family history of MI	cvd_riskfactors_rel.fammi
Family history of T1DM	cvd_riskfactors_rel.famidm
Family history of T2DM	cvd_riskfactors_rel.famniddm
Systolic blood pressure	cvd_riskfactors_rel.bps
Diastolic blood pressure	cvd_riskfactors_rel.bpd
Pulse pressure	cvd_riskfactors_rel.pulpres
Pulse rate	cvd_riskfactors_rel.pulse
Total cholesterol	cvd_riskfactors_rel.chl
Triglycerides	cvd_riskfactors_rel.trg

Table Variable	dataset.variable
HDLc	cvd_riskfactors_rel.hdl
LDLc	cvd_riskfactors_rel.ldl
Diabetes duration	cvd_riskfactors_rel.duration
C-peptide Diabetes duration < 5 years	cvd_riskfactors_rel.cpeplt5
C-peptide Diabetes duration ≥ 5 years	cvd_riskfactors_rel.cpepge5
AER	cvd_riskfactors_rel.caer
HbA1c (%)	cvd_riskfactors_rel.hba1c
HbA1c (mmol/mol)	nih_cvd30yr_1441.hbael_mmol

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

Characteristic	Overall Manuscript N=1,441	Overall DSIC N=1,441	Diff. N=0	Any-CVD No Manuscript n=1,257	Any-CVD No DSIC n=1,257	Diff. n=0
Design						
Treatment group (% conventional)	51	51	0	50	50	0
Cohort (% secondary)	50	50	0	48	48	0
Demographic						
Physical						
Sex (% men)	53	53	0	53	53	0
Age (years)	27 (22,32)	27 (22,32)	0 (0,0)	27 (21,32)	27 (21,32)	0 (0,0)
Adult vs. adolescent (< 18 years)	86	86	0	85	85	0
Weight men (kg)	74 (67,82)	74 (67,82)	0 (0,0)	74 (67,82)	74 (67,82)	0 (0,0)
Weight women (kg)	62 (56,69)	62 (56,69)	0 (0,0)	61 (55,68)	61 (55,68)	0 (0,0)
BMI men (kg/m ²)	24 (22,25)	24 (22,25)	0 (0,0)	23 (22,25)	23 (22,25)	0 (0,0)
BMI women (kg/m ²)	23 (21,25)	23 (21,25)	0 (0,0)	23 (21,25)	23 (21,25)	0 (0,0)
Behavioral						
Smoking (%)	19	19	0	17	17	0
Alcohol (% occasional or regular)	22	22	0	22	22	0
Exercise (% moderate or strenuous)	70	70	0	70	70	0
Family history (%)						
Hypertension	56	56	0	56	56	0
MI	49	49	0	47	47	0
T1DM	14	14	0	14	14	0

Characteristic	Overall Manuscript N=1,441	Overall DSIC N=1,441	Diff. N=0	Any-CVD No Manuscript n=1,257	Any-CVD No DSIC n=1,257	Diff. n=0
T2DM	9	9	0	9	9	0
Traditional						
Blood pressure						
Systolic (mmHg)	114 (106,122)	114 (106,122)	0 (0,0)	114 (106,120)	114 (106,120)	0 (0,0)
Diastolic (mmHg)	72 (68,80)	72 (68,80)	0 (0,0)	72 (68,80)	72 (68,80)	0 (0,0)
Pulse pressure (mmHg)	40 (34,48)	40 (34,48)	0 (0,0)	40 (34,48)	40 (34,48)	0 (0,0)
Pulse rate (bpm)	76 (68,84)	76 (68,84)	0 (0,0)	76 (68,82)	76 (68,82)	0 (0,0)
Lipids (mg/dL)						
Total cholesterol	174 (153,197)	174 (153,197)	0 (0,0)	172 (152,195)	172 (152,195)	0 (0,0)
Triglycerides	73 (55,94)	73 (55,94)	0 (0,0)	72 (55,93)	72 (55,93)	0 (0,0)
HDLc	49 (42,57)	49 (42,57)	0 (0,0)	50 (42,58)	50 (42,58)	0 (0,0)
LDLc	107 (91,127)	107 (91,127)	0 (0,0)	105 (89,125)	105 (89,125)	0 (0,0)
Diabetes-related						
History						
Diabetes duration (months)	49 (26,108)	49 (26,108)	0 (0,0)	47 (26,104)	47 (26,104)	0 (0,0)
C-peptide (nmol/L)						
Diabetes duration < 5 years	0.13 (0.04,0.25)	0.13 (0.04,0.25)	0 (0,0)	0.12 (0.04,0.24)	0.12 (0.04,0.24)	0 (0,0)
Diabetes duration ≥ 5 years	0.03 (0.03,0.04)	0.03 (0.03,0.04)	0 (0,0)	0.03 (0.03,0.04)	0.03 (0.03,0.04)	0 (0,0)
Nephropathy						
AER (mg/24 h)*	11 (7,19)	11 (7,19)	0 (0,0)	11 (7,17)	11 (7,17)	0 (0,0)
Glycemia						
HbA1c (%)**	8.8 (7.8,10.1)	8.8 (7.8,10.1)	0 (0,0)	8.8 (7.8,10.1)	8.8 (7.8,10.1)	0 (0,0)
HbA1c (mmol/mol)**	73 (62,87)	73 (62,87)	0 (0,0)	73 (62,87)	73 (62,87)	0 (0,0)

Characteristic	Any-CVD Yes Manuscript n=184	Any-CVD Yes DSIC n=184	Diff. n=0	HR Manuscript	HR DSIC	Diff.
Design						
Treatment group (% conventional)	55	55	0	1.29	1.29	0
Cohort (% secondary)	62	62	0	1.592	1.591	0.001
Demographic						
Physical						
Sex (% men)	52	52	0	1.004	1.004	0

Characteristic	Any-CVD Yes Manuscript n=184	Any-CVD Yes DSIC n=184	Diff. n=0	HR Manuscript	HR DSIC	Diff.
Age (years)	31 (27,35)	31 (27,35)	0 (0,0)	1.091	1.091	0
Adult vs. adolescent (< 18 years)	95	95	0	3.378	3.378	0
Weight men (kg)	76 (70,84)	76 (70,84)	0 (0,0)	1.011	1.011	0
Weight women (kg)	65 (57,70)	65 (57,70)	0 (0,0)	1.034	1.034	0
BMI men (kg/m ²)	24 (22,26)	24 (22,26)	0 (0,0)	1.065	1.065	0
BMI women (kg/m ²)	24 (22,26)	24 (22,26)	0 (0,0)	1.096	1.096	0
Behavioral						
Smoking (%)	27	27	0	1.834	1.834	0
Alcohol (% occasional or regular)	20	20	0	0.888	0.888	0
Exercise (% moderate or strenuous)	71	71	0	1.067	1.067	0
Family history (%)						
Hypertension	59	59	0	1.123	1.123	0
MI	60	60	0	1.615	1.615	0
T1DM	17	17	0	1.294	1.294	0
T2DM	13	13	0	1.592	1.592	0
Traditional						
Blood pressure						
Systolic (mmHg)	116 (110,124)	116 (110,124)	0 (0,0)	1.016	1.016	0
Diastolic (mmHg)	74 (68,80)	74 (68,80)	0 (0,0)	1.021	1.021	0
Pulse pressure (mmHg)	42 (34,48)	42 (34,48)	0 (0,0)	1.006	1.006	0
Pulse rate (bpm)	78 (72,88)	78 (72,88)	0 (0,0)	1.019	1.018	0.001
Lipids (mg/dL)						
Total cholesterol	185 (162,211)	185 (162,211)	0 (0,0)	1.012	1.012	0
Triglycerides	80 (59,97)	80 (58,97)	0 (1,0)	1.614	1.614	0
HDLc	47 (41,55)	47 (41,55)	0 (0,0)	0.992	0.992	0
LDLc	119 (98,144)	119 (98,144)	0 (0,0)	1.014	1.014	0
Diabetes-related						
History						
Diabetes duration (months)	69 (29,130)	69 (29,130)	0 (0,0)	1.004	1.004	0
C-peptide (nmol/L)						
Diabetes duration < 5 years	0.13 (0.03,0.29)	0.13 (0.03,0.29)	0 (0,0)	1.5571	1.557	0

Characteristic	Any-CVD Yes Manuscript n=184	Any-CVD Yes DSIC n=184	Diff. n=0	HR Manuscript	HR DSIC	Diff.
Diabetes duration ≥ 5 years	0.03 (0.03,0.03)	0.03 (0.03,0.03)	0 (0,0)	0.079	0.079	0
Nephropathy						
AER (mg/24 h)*	13 (7,20)	13 (7,20)	0 (0,0)	1.17	1.17	0
Glycemia						
HbA1c (%)**	9.0 (8.1, 10.3)	9.0 (8.1,10.3)	0 (0,0)	1.093	1.093	0
HbA1c (mmol/mol)**	75 (65,89)	75 (65,90)	0 (0,1)	1.085	1.008	0.077

Characteristic	95% CI Manuscript	95% CI DSIC	Diff.
Design			
Treatment group (% conventional)	0.964, 1.725	0.964, 1.725	0, 0
Cohort (% secondary)	1.18, 2.147	1.18, 2.147	0, 0
Demographic			
Physical			
Sex (% men)	0.752, 1.341	0.752, 1.341	0, 0
Age (years)	1.066, 1.116	1.066, 1.116	0, 0
Adult vs. adolescent (< 18 years)	1.727, 6.611	1.727, 6.611	0, 0
Weight men (kg)	0.993, 1.029	0.993, 1.029	0, 0
Weight women (kg)	1.011, 1.057	1.011, 1.057	0, 0
BMI men (kg/m ²)	0.993, 1.142	0.993, 1.142	0, 0
BMI women (kg/m ²)	1.022, 1.175	1.022, 1.175	0, 0
Behavioral			
Smoking (%)	1.325, 2.538	1.325, 2.538	0, 0
Alcohol (% occasional or regular)	0.617, 1.279	0.617, 1.279	0, 0
Exercise (% moderate or strenuous)	0.775, 1.468	0.775, 1.468	0, 0
Family history (%)			
Hypertension	0.838, 1.506	0.838, 1.506	0, 0
MI	1.202, 2.169	1.202, 2.169	0, 0
T1DM	0.884, 1.894	0.884, 1.894	0, 0
T2DM	1.037, 2.446	1.037, 2.446	0, 0
Traditional			

Characteristic	95% CI Manuscript	95% CI DSIC	Diff.
Blood pressure			
Systolic (mmHg)	1.004, 1.029	1.004, 1.029	0, 0
Diastolic (mmHg)	1.004, 1.038	1.004, 1.038	0, 0
Pulse pressure (mmHg)	0.991, 1.02	0.991, 1.02	0,0
Pulse rate (bpm)	1.006, 1.031	1.006, 1.03	0, 0.001
Lipids (mg/dL)			
Total cholesterol	1.008, 1.016	1.008, 1.016	0, 0
Triglycerides	1.184, 2.199	1.185, 2.2	0.001, 0.001
HDLc	0.98, 1.004	0.98, 1.004	0, 0
LDLc	1.01, 1.019	1.01, 1.019	0, 0
Diabetes-related			
History			
Diabetes duration (months)	1.001, 1.007	1.001, 1.007	0, 0
C-peptide (nmol/L)			
Diabetes duration < 5 years	0.325, 7.458	0.325, 7.458	0,0
Diabetes duration ≥ 5 years	0, 33.708	0, 33.759	0, 0.051
Nephropathy			
AER (mg/24 h)*	0.978, 1.401	0.978, 1.401	0, 0
Glycemia			
HbA1c (%)**	1.003, 1.19	1.003, 1.19	0, 0
HbA1c (mmol/mol)**	1.003, 1.173	1, 1.016	0.003, 0.157

*The DCC noted that the values presented for the median and quartiles for AER in the manuscript are incorrect; the values in this table are the correct values.

** Note that the values published for hazard ratio and 95% Confidence Interval for HbA1c (mmol/mol) are actually those for HbA1c (%). This table reflects the correct values.

Attachment A: SAS Code

```
*** EDIC CVD Data DSIC;
*** CVD risk factors data;
*** Programmer: Allyson Mateja;
*** Date: 1/4/2016;
*** Modified 2/17/17;

title 'DCCT/EDIC CVD Risk Factors Data DSIC';
title2 ' ';

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

libname cvddata '/prj/niddk/ims_analysis/DCCT_EDIC/private_orig_data/CVD_Data/';
libname riskdata '/prj/niddk/ims_analysis/DCCT_EDIC/private_orig_data/CVD_RISK/';

data cvd_events;
    set cvddata.nih_cvd30yr_ev_366;

proc contents data = cvd_events;

proc sort data = cvd_events nodupkey out=mask_pat_only;
    by mask_pat;

data cvd_risk_factors;
    set riskdata.cvd_riskfactors_rel;

proc contents data = cvd_risk_factors;

data cvd_all_subjects;
    set cvddata.nih_cvd30yr_1441;

proc freq data = cvd_risk_factors;
    tables year;

data baseline_data;
    set cvd_risk_factors;
    if year=0;

proc sort data = baseline_data;
    by mask_pat;

proc sort data = cvd_all_subjects;
    by mask_pat;

data baseline_data;
    merge baseline_data (in=val1)
           mask_pat_only (in=val2 keep=mask_pat)
```

```

        cvd_all_subjects (keep=mask_pat carvtime1 aer00 HBAEL_MMOL trg00 carv hard hardtime1);
    by mask_pat;
    if val2 then any_cvd = 1;
    else any_cvd=0;
    if drinks = 1 then drinking = 0;
    else if drinks in (2,3) then drinking = 1;
    if exercise in (1, .) then exercising = 0;
    else if exercise in (2,3) then exercising = 1;
    transtrg = log(trg);
    transaer = log(caer);
    HBAMMOL = ((hbalc - 2.152) * 10.931);
    if vall then output;

proc freq data = baseline_data;
    tables any_cvd;
    format any_cvd yesnof.;
    title3 'Table 1 - Any CVD';

proc sort data = baseline_data;
    by any_cvd;

proc freq data = baseline_data;
    tables group retbase;
    title3 'Table 1 - Design';

proc freq data = baseline_data;
    tables group retbase;
    by any_cvd;
    format any_cvd yesnof.;

proc phreg data=baseline_data;
    class group (ref='EXPERIMENTAL');
    model carvtime1*any_cvd(0) = group /risklimits;
    title3 'Table 1 - Treatment group hazard ratio';

proc phreg data=baseline_data;
    class retbase (ref='PRIM');
    model carvtime1*any_cvd(0) = retbase /risklimits;
    title3 'Table 1 - Cohort hazard ratio';

proc freq data = baseline_data;
    tables sex;
    title3 'Table 1 - Sex';

proc freq data = baseline_data;
    tables sex;
    by any_cvd;
    format any_cvd yesnof.;

proc phreg data=baseline_data;
    class sex (ref='F');

```

```

        model carvtime1*any_cvd(0) = sex /risklimits;
        title3 'Table 1 - Sex hazard ratio';

proc means data = baseline_data n median p25 p75;
    var age;
    class any_cvd;
    types () any_cvd;
    format any_cvd yesnof.;
    title3 'Table 1 - Age';

proc phreg data=baseline_data;
    model carvtime1*any_cvd(0) = age /risklimits;
    title3 'Table 1 - Age hazard ratio';

proc freq data = baseline_data;
    tables adult;
    title3 'Table 1 - Adult';

proc freq data = baseline_data;
    tables adult;
    by any_cvd;
    format any_cvd yesnof.;

proc phreg data=baseline_data;
    class adult (ref='0');
    model carvtime1*any_cvd(0) = adult /risklimits;
    title3 'Table 1 - Adult hazard ratio';

proc means data = baseline_data n median p25 p75;
    var weight;
    class any_cvd;
    types () any_cvd;
    format any_cvd yesnof.;
    where sex = 'M';
    title3 'Table 1 - Weight men';

proc phreg data=baseline_data;
    model carvtime1*any_cvd(0) = weight /risklimits;
    where sex = 'M';
    title3 'Table 1 - Weight (men) hazard ratio';

proc means data = baseline_data n median p25 p75;
    var weight;
    class any_cvd;
    types () any_cvd;
    format any_cvd yesnof.;
    where sex = 'F';
    title3 'Table 1 - Weight women';

proc phreg data=baseline_data;
    model carvtime1*any_cvd(0) = weight /risklimits;

```

```

        where sex = 'F';
        title3 'Table 1 - Weight (women) hazard ratio';

proc means data = baseline_data n median p25 p75;
    var bmi;
    class any_cvd;
    types () any_cvd;
    format any_cvd yesnof.;
    where sex = 'M';
    title3 'Table 1 - BMI men';

proc phreg data=baseline_data;
    model carvtimel*any_cvd(0) = bmi /risklimits;
    where sex = 'M';
    title3 'Table 1 - BMI (men) hazard ratio';

proc means data = baseline_data n median p25 p75;
    var bmi;
    class any_cvd;
    types () any_cvd;
    format any_cvd yesnof.;
    where sex = 'F';
    title3 'Table 1 - BMI women';

proc phreg data=baseline_data;
    model carvtimel*any_cvd(0) = bmi /risklimits;
    where sex = 'F';
    title3 'Table 1 - BMI (women) hazard ratio';

proc freq data = baseline_data;
    tables smokes;
    title3 'Table 1 - Smoking';

proc freq data = baseline_data;
    tables smokes;
    by any_cvd;
    format any_cvd yesnof.;

proc phreg data=baseline_data;
    class smokes (ref='0');
    model carvtimel*any_cvd(0) = smokes /risklimits;
    title3 'Table 1 - Smoking hazard ratio';

proc freq data = baseline_data;
    tables drinking;
    title3 'Table 1 - Alcohol';

proc freq data = baseline_data;
    tables drinking;
    by any_cvd;
    format any_cvd yesnof.;

```

```

proc phreg data=baseline_data;
  class drinking (ref='0');
  model carvtimel*any_cvd(0) = drinking /risklimits;
  title3 'Table 1 - Drinking hazard ratio';

proc freq data = baseline_data;
  tables exercising;
  title3 'Table 1 - Exercise';

proc freq data = baseline_data;
  tables exercising;
  by any_cvd;
  format any_cvd yesnof.;

proc phreg data=baseline_data;
  class exercising (ref='0');
  model carvtimel*any_cvd(0) = exercising /risklimits;
  title3 'Table 1 - Exercise hazard ratio';

proc freq data = baseline_data;
  tables famht;
  title3 'Table 1 - Family history of hypertension';

proc freq data = baseline_data;
  tables famht;
  by any_cvd;
  format any_cvd yesnof.;

proc phreg data=baseline_data;
  class famht (ref='0');
  model carvtimel*any_cvd(0) = famht /risklimits;
  title3 'Table 1 - Family history of hypertension hazard ratio';

proc freq data = baseline_data;
  tables fammi;
  title3 'Table 1 - Family history of MI';

proc freq data = baseline_data;
  tables fammi;
  by any_cvd;
  format any_cvd yesnof.;

proc phreg data=baseline_data;
  class fammi (ref='0');
  model carvtimel*any_cvd(0) = fammi /risklimits;
  title3 'Table 1 - Family history of MI hazard ratio';

proc freq data = baseline_data;
  tables famiddm;
  title3 'Table 1 - Family history of T1DM';

```



```

proc freq data = baseline_data;
  tables famiddm;
  by any_cvd;
  format any_cvd yesnof.;

proc phreg data=baseline_data;
  class famiddm (ref='0');
  model carvtime1*any_cvd(0) = famiddm /risklimits;
  title3 'Table 1 - Family history of T1DM hazard ratio';

proc freq data = baseline_data;
  tables famniddm;
  title3 'Table 1 - Family history of T2DM';

proc freq data = baseline_data;
  tables famniddm;
  by any_cvd;
  format any_cvd yesnof.;

proc phreg data=baseline_data;
  class famniddm (ref='0');
  model carvtime1*any_cvd(0) = famniddm /risklimits;
  title3 'Table 1 - Family history of T2DM hazard ratio';

proc means data = baseline_data n median p25 p75;
  var bps;
  class any_cvd;
  types () any_cvd;
  format any_cvd yesnof.;
  title3 'Table 1 - Systolic Blood Pressure';

proc phreg data=baseline_data;
  model carvtime1*any_cvd(0) = bps /risklimits;
  title3 'Table 1 - Systolic blood pressure hazard ratio';

proc means data = baseline_data n median p25 p75;
  var bpd;
  class any_cvd;
  types () any_cvd;
  format any_cvd yesnof.;
  title3 'Table 1 - Diastolic Blood Pressure';

proc phreg data=baseline_data;
  model carvtime1*any_cvd(0) = bpd /risklimits;
  title3 'Table 1 - Diastolic blood pressure hazard ratio';

proc means data = baseline_data n median p25 p75;
  var pulpres;
  class any_cvd;
  types () any_cvd;

```

```

format any_cvd yesnof.;
title3 'Table 1 - Pulse Pressure';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = pulpres /risklimits;
title3 'Table 1 - Pulse pressure hazard ratio';

proc means data = baseline_data n median p25 p75;
var pulse;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - Pulse Rate';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = pulse /risklimits;
title3 'Table 1 - Pulse rate hazard ratio';

proc means data = baseline_data n median p25 p75;
var chl;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - Total Cholesterol';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = chl /risklimits;
title3 'Table 1 - Total cholesterol hazard ratio';

proc means data = baseline_data n median p25 p75 noprint;
var transtrg;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
output out = mean_data mean=mean p25=p25 p75=p75;
title3 'Table 1 - Triglycerides';

data mean_data;
set mean_data;
mean=exp(mean);
p25=exp(p25);
p75=exp(p75);

proc print data = mean_data;

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = transtrg /risklimits;
title3 'Table 1 - Triglycerides hazard ratio';

proc means data = baseline_data n median p25 p75;
var hdl;

```

```

class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - HDLc';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = hdl /risklimits;
title3 'Table 1 - HDLc hazard ratio';

proc means data = baseline_data n median p25 p75;
var ldl;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - LDLc';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = ldl /risklimits;
title3 'Table 1 - LDLc hazard ratio';

proc means data = baseline_data n median p25 p75;
var duration;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - Diabetes duration';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = duration /risklimits;
title3 'Table 1 - Diabetes duration hazard ratio';

proc means data = baseline_data n median p25 p75;
var cpeplt5;
where cpeplt5 ne 0;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - C-peptide diabetes duration < 5 years';

proc phreg data=baseline_data;
model carvtimel*any_cvd(0) = cpeplt5 /risklimits;
where cpeplt5 ne 0;
title3 'Table 1 - C-peptide diabetes duration < 5 years hazard ratio';

proc means data = baseline_data n median p25 p75;
var cpepge5;
where cpepge5 ne 0;
class any_cvd;
types () any_cvd;
format any_cvd yesnof.;
title3 'Table 1 - C-peptide diabetes duration >= 5 years';

```

```

proc phreg data=baseline_data;
  model carvtimel*any_cvd(0) = cpepge5 /risklimits;
  where cpepge5 ne 0;
  title3 'Table 1 - C-peptide diabetes duration >= 5 years hazard ratio';

proc means data = baseline_data n median p25 p75 noprint;
  var transaer;
  class any_cvd;
  types () any_cvd;
  format any_cvd yesnof.;
  output out = mean_data mean=mean p25=p25 p75=p75;
  title3 'Table 1 - AER';

data mean_data;
  set mean_data;
  mean=exp(mean);
  p25=exp(p25);
  p75=exp(p75);

proc print data = mean_data;

proc phreg data=baseline_data;
  model carvtimel*any_cvd(0) = transaer /risklimits;
  title3 'Table 1 - AER hazard ratio';

proc means data = baseline_data n median p25 p75;
  var hbalc;
  class any_cvd;
  types () any_cvd;
  format any_cvd yesnof.;
  title3 'Table 1 - HbA1c (%)';

proc phreg data=baseline_data;
  model carvtimel*any_cvd(0) = hbalc /risklimits;
  title3 'Table 1 - HbA1c (%) hazard ratio';

proc means data = baseline_data n median p25 p75;
  var hbammol;
  class any_cvd;
  types () any_cvd;
  format any_cvd yesnof.;
  title3 'Table 1 - HbA1c (mmol/mol)';

proc phreg data=baseline_data;
  model carvtimel*any_cvd(0) = hbammol /risklimits;
  title3 'Table 1 - HbA1c (mmol/mol) hazard ratio';

ods graphics on;

PROC LIFETEST DATA=baseline_data plots=survival(cl atrisk=0 to 30 by 5 nocensor) alpha=0.05 outsurv=out_data_all;

```

```

TIME carvtime1 * carv (0);
RUN;

PROC LIFETEST DATA=baseline_data plots=survival(cl atrisk=0 to 30 by 5 nocensor) alpha=0.05 outsurv=out_data_mace;
TIME hardtime1 * hard (0);
RUN;

data out_data;
    merge out_data_all (in=val1 rename = (carvtime1 = years survival = any_cvd_survival sdf_lcl = sdf_lcl_any_cvd sdf_ucl =
sdf_ucl_any_cvd))
        out_data_mace (in=val2 rename = (hardtime1 = years survival = mace_survival sdf_lcl = sdf_lcl_mace sdf_ucl =
sdf_ucl_mace));
    by years;
    if val1 or val2 then output;

ods csv file = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/dsic_cvd_risk_factors.csv';
proc print data = out_data noobs;
run;
ods csv close;

```