Dataset Integrity Check for the DCCT/EDIC Microvascular Complications CVD Dataset

Prepared by NIDDK-CR
August 16, 2021

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 Epidemiology of Diabetes Interventions and Complications (EDIC) study was initiated as follow-up to examine the long-term effects of the original Diabetes Control and Complications Trial (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 (CVD) were measured at different intervals for participants. The association between microvascular complications and subsequent risk of CVD was assessed using Cox proportional hazard models.

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the DCCT/EDIC folder in the data package. For this replication, variables were taken from the “micro_cvd.sas7bdat” dataset.

4 Statistical Methods

Analyses were performed to replicate results for the data published by Gubitosi-Klug et al. [1] for Associations of Microvascular Complications With the Risk of Cardiovascular Disease in Type 1 Diabetes. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline characteristics of the DCCT/EDIC participants by status of CVD or MACE over the combined DCCT/EDIC follow-up, Table A lists the variables that were used in the replication, and Tables B1-B2 compares the results calculated from the archived data files to the results published in Table 1. The results of the replication are a match to the published results.

6 Conclusions

The NIDDK Central Repository is confident that the DCCT/EDIC Microvascular Complications CVD data files to be distributed are a true copy of the study data.

7 References

[1] Gubitosi-Klug R, Gao X, Pop-Busui R, de Boer IH, White N, Aiello LP, Miller R, Palmer J, Tamborlane W, Wallia A, Kosiborod M, Lachin JM, Bebu I. Associations of Microvascular Complications With the Risk of Cardiovascular Disease in Type 1 Diabetes. Diabetes Care, 44(7), 1499-1505, July 2021. doi: https://doi.org/10.2337/dc20-3104

Table A: Variables used to replicate Table 1 – Baseline characteristics of the DCCT/EDIC participants by status of CVD or MACE over the combined DCCT/EDIC follow-up

Table Variable	**dataset.variable**
Age (years)	micro_cvd.age
Sex (% female)	micro_cvd.female
T1D duration (months)	micro_cvd.duration
Family history of MI (%)	micro_cvd.fammi
Smoking (%)	micro_cvd.fsmokes
Systolic blood pressure (mmHg)	micro_cvd.basebps
Pulse (bpm)	micro_cvd.basepulse
LDL (mg/dL)	micro_cvd.baseldl
Triglycerides (mg/dL)	micro_cvd.basetrg
HbA_{1C} (%)	micro_cvd.basehba1c
AER (mg/24 h)	micro_cvd.basecaer
eGFR (mL/min/1.73 m^2)	micro_cvd.baseckd_gfr

Table B1: Comparison of values computed in integrity check to reference article Table 1 values (Overall and Any CVD)

Variable	Manuscript Overall (n=1441)	DSIC Overall (n=1441)	Diff. (n=0)	Manuscript No CVD (n=1202)	DSIC No CVD (n=1202)	Diff. (n=0)	Manuscript With CVD (n=239)	DSIC With CVD (n=239)	Diff. (n=0)
Age (years)	26.8 (7.1)	26.8 (7.1)	0 (0)	26.2 (7.1)	26.2 (7.1)	0 (0)	29.9 (6.3)	29.9 (6.3)	0 (0)
Sex (% female)	47.2	47.2	0	47.8	47.8	0	43.9	43.9	0
T1D duration (months)	69.8 (49.7)	69.8 (49.7)	0 (0)	67.9 (49.0)	67.9 (49.0)	0 (0)	79.4 (51.9)	79.4 (51.9)	0 (0)
Family history of MI (%)	48.9	48.9	0	46.8	46.8	0	59.0	59.0	0
Smoking (%)	18.5	18.5	0	16.6	16.6	0	28.0	28.0	0
Systolic blood pressure (mmHg)	114.5 (11.4)	114.5 (11.4)	0 (0)	114.2 (11.3)	114.2 (11.3)	0 (0)	116.2 (11.4)	116.2 (11.4)	0 (0)
Pulse (bpm)	76.1 (11.1)	76.1 (11.1)	0 (0)	75.8 (11.1)	75.8 (11.1)	0 (0)	77.6 (11.1)	77.6 (11.1)	0 (0)
LDL (mg/dL)	109.7 (29.1)	109.7 (29.1)	0 (0)	107.9 (28.6)	107.9 (28.6)	0 (0)	118.8 (30.0)	118.8 (30.0)	0 (0)
Triglycerides (mg/dL)	81.3 (47.5)	81.3 (47.5)	0 (0)	79.9 (46.7)	79.9 (46.7)	0 (0)	88.6 (50.7)	88.6 (50.7)	0 (0)
HbA_{1C} (%)	8.9 (1.6)	8.9 (1.6)	0 (0)	8.9 (1.6)	8.9 (1.6)	0 (0)	9.1 (1.6)	9.1 (1.6)	0 (0)
HbA_{1C} (mmol/mol)[1]	74 (17.4)	73.6 (17.2)	0.4 (0.2)	74 (17.4)	73.2 (17.3)	0.8 (0.1)	76 (17.5)	76 (17.9)	0 (0.4)
AER (mg/24 h)	15.9 (18.8)	15.9 (18.8)	0 (0)	15.7 (18.7)	15.7 (18.7)	0 (0)	17.0 (19.0)	17.0 (19.0)	0 (0)
eGFR (mL/min/1.73 m^2)	126.1 (14.2)	126.1 (14.2)	0 (0)	126.6 (14.3)	126.6 (14.3)	0 (0)	123.9 (13.5)	123.9 (13.5)	0 (0)

[1] HbA_{1c} (mmol/mol) calculated using formula: 10.929 * (A1C(%) - 2.15)

Table B2: Comparison of values computed in integrity check to reference article Table 1 values (MACE)

Variable	Manuscript No MACE (n=1321)	DSIC No MACE (n=1321)	Diff. (n=0)	Manuscript With MACE (n=120)	DSIC With MACE (n=120)	Diff. (n=0)
Age (years)	26.5 (7.1)	26.5 (7.1)	0 (0)	30.3 (6.1)	30.3 (6.1)	0 (0)
Sex (% female)	47.9	47.9	0	39.2	39.2	0
T1D duration (months)	68.6 (49.1)	68.6 (49.1)	0 (0)	82.7 (54.3)	82.7 (54.3)	0 (0)
Family history of MI (%)	48.4	48.4	0	53.3	53.3	0
Smoking (%)	17.1	17.1	0	33.3	33.3	0
Systolic blood pressure (mmHg)	114.3 (11.3)	114.3 (11.3)	0 (0)	116.7 (11.3)	116.7 (11.3)	0 (0)
Pulse (bpm)	75.9 (11.1)	75.9 (11.1)	0 (0)	77.9 (11.2)	77.9 (11.2)	0 (0)
LDL (mg/dL)	109.0 (28.9)	109.0 (28.9)	0 (0)	117.5 (30.2)	117.5 (30.2)	0 (0)
Triglycerides (mg/dL)	80.1 (46.9)	80.1 (46.9)	0 (0)	94.8 (51.6)	94.8 (51.6)	0 (0)
HbA_{1C} (%)	8.9 (1.6)	8.9 (1.6)	0 (0)	9.2 (1.7)	9.2 (1.7)	0 (0)
HbA_{1C} (mmol/mol)[1]	74 (17.5)	73.4 (17.3)	0.6 (0.2)	77 (18.6)	77 (18.4)	0 (0.2)
AER (mg/24 h)	15.8 (18.5)	15.8 (18.5)	0 (0)	17.8 (21.7)	17.8 (21.7)	0 (0)
eGFR (mL/min/1.73 m^2)	126.4 (14.2)	126.4 (14.2)	0 (0)	122.7 (14.0)	122.7 (14.0)	0 (0)

[1]HbA_{1c} (mmol/mol) calculated using formula: 10.929 * (A1C(%) - 2.15)

Attachment A: SAS Code

```
libname cvd "X:\NIDDK\niddk-
dr_studies1\DCCT_EDIC\private_orig_data\Microvascular CVD";

*************************************************************************
****;
*Replicating Table 1 from Associations of Microvascular Complications
*;
*With the Risk of Cardiovascular Disease in Type1 Diabetes -Gubitosi-
Klug*;
*et. al. 2021
*;
*************************************************************************
****;

*temp dataset;
data cvd; set cvd.micro_cvd;
run;

proc contents data=cvd;
run;

proc freq data=cvd;
tables dtedyear;
run;

******************;
*Table 1 - OVERALL;
******************;

*AGE;
proc means data=cvd mean std;
var age;
where dtedyear = 0;
run;

*SEX;
proc freq data=cvd;
tables female;
where dtedyear = 0;
run;

*T1D DURATION;
proc means data=cvd mean std;
var duration;
where dtedyear = 0;
run;

*FAMILY HISTORY MI;
```

```
proc freq data=cvd;
tables fammi;
where dtedyear = 0;
run;

*SMOKING;
proc freq data=cvd;
tables fsmokes;
where dtedyear = 0;
run;

*SYSTOLIC BLOOS PRESSURE;
proc means data=cvd mean std;
var basebps;
where dtedyear = 0;
run;

*PULSE;
proc means data=cvd mean std;
var basepulse;
where dtedyear = 0;
run;

*LDL;
proc means data=cvd mean std;
var baseldl;
where dtedyear = 0;
run;

*TRIGLYCERIDES;
proc means data=cvd mean std;
var basetrg;
where dtedyear = 0;
run;

*HbA1c;
proc means data=cvd mean std;
var basehba1c;
where dtedyear = 0;
run;

*HbA1c conversion to mmol/mol (A1C(mmol/mol) = 10.929 * (A1C(%) -
2.15));
data one; set cvd;
a1c_calc = 10.929*(basehba1c - 2.15);
run;

proc means data=one mean std;
var a1c_calc;
run;

*AER;
```

```
proc means data=cvd mean std;
var basecaer;
where dtedyear = 0;
run;

*eGFR;
proc means data=cvd mean std;
var baseckd_gfr;
where dtedyear = 0;
run;


*******************************;
*      Table 1 - ANY CVD       *;
*******************************;
data any_cvd; set cvd;
where keepcarv = 1;
run;

*AGE;
proc means data=any_cvd mean std;
var age;
class evercarv;
where dtedyear = 0;
run;

*SEX;
proc freq data=any_cvd;
tables female*evercarv;
where dtedyear = 0;
run;

*T1D DURATION;
proc means data=any_cvd mean std;
var duration;
class evercarv;
where dtedyear = 0;
run;

*FAMILY HISTORY OF MI;
proc freq data=any_cvd;
tables fammi*evercarv/norow nopercent;
where dtedyear = 0;
run;

*SMOKING;
proc freq data=any_cvd;
tables Fsmokes*evercarv / norow nopercent;
where dtedyear = 0;
run;

*SYSTOLIC BLOOD PRESSURE;
```

```
proc means data=any_cvd mean std;
var basebps;
class evercarv;
where dtedyear = 0;
run;

*PULSE;
proc means data=any_cvd mean std;
var basepulse;
class evercarv;
where dtedyear = 0;
run;

*LDL;
proc means data=any_cvd mean std;
var baseldl;
class evercarv;
where dtedyear = 0;
run;

*TRIGLYCERIDES;
proc means data=any_cvd mean std;
var basetrg;
class evercarv;
where dtedyear = 0;
run;

*HBA1C %;
proc means data=any_cvd mean std;
var basehba1c;
class evercarv;
where dtedyear = 0;
run;

*HbA1c conversion to mmol/mol (A1C(mmol/mol) = 10.929 * (A1C(%) -
2.15));
data two; set any_cvd;
a1c_calc = 10.929*(basehba1c - 2.15);
run;

proc means data=two mean std;
var a1c_calc;
class evercarv;
where dtedyear = 0;
run;

*AER;
proc means data=any_cvd mean std;
var basecaer;
class evercarv;
where dtedyear = 0;
run;
```

```
*eGFR;
proc means data=any_cvd mean std;
var baseckd_gfr;
class evercarv;
where dtedyear =0;
run;

*****************************;
*           MACE              *;
*****************************;

data mace; set cvd;
if keephard =1;
run;

*AGE;
proc means data=mace mean std;
var age;
class everhard;
where dtedyear = 0;
run;

*SEX;
proc freq data=mace;
tables female*everhard/ norow nopercent;
where dtedyear = 0;
run;

*T1D DURATION;
proc means data=mace mean std;
var duration;
class everhard;
where dtedyear = 0;
run;

*FAMILY HISTORY MI;
proc freq data=mace;
tables fammi*everhard/ norow nopercent;
where dtedyear = 0;
run;

*SMOKING;
proc freq data=mace;
tables fsmokes*everhard/ norow nopercent;
where dtedyear = 0;
run;

*SYTOLIC BLOOD PRESSURE;
proc means data=mace mean std;
var basebps;
class everhard;
```

```
where dtedyear = 0;
run;

*PULSE;
proc means data=mace mean std;
var basepulse;
class everhard;
where dtedyear = 0;
run;

*LDL;
proc means data=mace mean std;
var baseldl;
class everhard;
where dtedyear = 0;
run;

*TRIGLYCERIDES;
proc means data=mace mean std;
var basetrg;
class everhard;
where dtedyear = 0;
run;

*HBA1C %;
proc means data=mace mean std;
var basehba1c;
class everhard;
where dtedyear = 0;
run;

*HbA1c conversion to mmol/mol (A1C(mmol/mol) = 10.929 * (A1C(%) -
2.15));
data three; set mace;
a1c_calc = 10.929*(basehba1c - 2.15);
run;

proc means data=three mean std;
var a1c_calc;
class everhard;
where dtedyear = 0;
run;

*AER;
proc means data=mace mean std;
var basecaer;
class everhard;
where dtedyear = 0;
run;

*eGFR;
proc means data=mace mean std;
```

```
var baseckd_gfr;
class everhard;
where dtedyear = 0;
run;
```