Dataset Integrity Check for DCCT/EDIC CVD 30 Year Update Data Files

Prepared by Allyson Mateja
IMS, Inc.
3901 Calverton Blvd, Suite 200, Calverton, MD 20705
January 5, 2017

Contents

1 Standard Disclaimer

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

2 Study Background

The 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 DDCT 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 and nih_cvd30yr_1441.sas7bdat.

4 Statistical Methods

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

5 Results

For Table 1 in the publication [1], Cardiovascular events in each original treatment group of the DCCT, Table A lists the variables that can be 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 an exact match to the results in publication [1].

For Figure 1 in the publication [1], Table C 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].

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] Gubitosi-Klug, R., et al.; DCCT/EDIC Research Group. Intensive Diabetes Treatment and Cardiovascular Outcomes in type 1 Diabetes: The DCCT/EDIC Study 30-Year Follow-up. Diabetes Care 2016;39:686-693.

Table A: Variables used to replicate Table 1: Cardiovascular events in each original treatment group of the DCCT

Table Variable	**dataset.variable**
Treatment group	nih_cvd30yr_1441.group
CVD Event	nih_cvd30yr_ev_366.cvd_event
MACE	nih_cvd30yr_1441.hard
Initial event	nih_cvd30yr_ev_366.init_event

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

Intensive-treatment group:

Event	Patients n (%) Manuscript	Patients n (%) DSIC	Diff.	Events n Manuscript	Events n DSIC	Diff.
Any CVD Events	82 (11.5)	82 (11.5)	0 (0)	149	149	0
1. Nonfatal acute MI	24 (3.4)	24 (3.4)	0 (0)	26	26	0
2. Nonfatal cerebrovascular event	8 (1.1)	8 (1.1)	0 (0)	9	9	0
3. Death from CVD	9 (1.3)	9 (1.3)	0 (0)	9	9	0
4. Silent MI	20 (2.8)	20 (2.8)	0 (0)	21	21	0
5. Confirmed angina	20 (2.8)	20 (2.8)	0 (0)	20	20	0
6. Revascularization	40 (5.6)	40 (5.6)	0 (0)	62	62	0
7. CHF	2 (0.3)	2 (0.3)	0 (0)	2	2	0
MACE	39 (5.5)	39 (5.5)	0 (0)	44	44	0
1. Nonfatal acute MI	24 (3.4)	24 (3.4)	0 (0)	26	26	0
2. Nonfatal cerebrovascular event	8 (1.1)	8 (1.1)	0 (0)	9	9	0
3. Death from CVD	9 (1.3)	9 (1.3)	0 (0)	9	9	0

Event	Initial events n Manuscript	Initial events n DSIC	Diff.	Secondary events n Manuscript	Secondary events n DSIC	Diff.
Any CVD Events	82	82	0	67	67	0
1. Nonfatal acute MI	19	19	0	7	7	0
2. Nonfatal cerebrovascular event	7	7	0	2	2	0
3. Death from CVD	5	5	0	4	4	0
4. Silent MI	19	19	0	2	2	0

Event	Initial events n Manuscript	Initial events n DSIC	Diff.	Secondary events n Manuscript	Secondary events n DSIC	Diff.
5. Confirmed angina	11	11	0	9	9	0
6. Revascularization	20	20	0	42	42	0
7. CHF	1	1	0	1	1	0
MACE	39	39	0	5	5	0
1. Nonfatal acute MI	24	24	0	2	2	0
2. Nonfatal cerebrovascular event	8	8	0	1	1	0
3. Death from CVD	7	7	0	2	2	0

Conventional-treatment group

Event	Patients n (%) Manuscript	Patients n (%) DSIC	Diff.	Events n Manuscript	Events n DSIC	Diff.
Any CVD Events	102 (14.0)	102 (14.0)	0 (0)	217	217	0
1. Nonfatal acute MI	29 (4.0)	29 (4.0)	0 (0)	35	35	0
2. Nonfatal cerebrovascular event	12 (1.6)	12 (1.6)	0 (0)	13	13	0
3. Death from CVD	16 (2.2)	16 (2.2)	0 (0)	16	16	0
4. Silent MI	33 (4.5)	33 (4.5)	0 (0)	36	36	0
5. Confirmed angina	22 (3.0)	22 (3.0)	0 (0)	33	33	0
6. Revascularization	48 (6.6)	48 (6.6)	0 (0)	71	71	0
7. CHF	10 (1.4)	10 (1.4)	0 (0)	13	13	0
MACE	49 (6.7)	49 (6.7)	0 (0)	64	64	0
1. Nonfatal acute MI	29 (4.0)	29 (4.0)	0 (0)	35	35	0
2. Nonfatal cerebrovascular event	12 (1.6)	12 (1.6)	0 (0)	13	13	0
3. Death from CVD	16 (2.2)	16 (2.2)	0 (0)	16	16	0

Event	Initial events n Manuscript	Initial events n DSIC	Diff.	Secondary events n Manuscript	Secondary events n DSIC	Diff.
Any CVD Events	102	102	0	115	115	0
1. Nonfatal acute MI	23	23	0	12	12	0
2. Nonfatal cerebrovascular event	11	11	0	2	2	0
3. Death from CVD	8	8	0	8	8	0
4. Silent MI	24	24	0	12	12	0
5. Confirmed angina	11	11	0	22	22	0

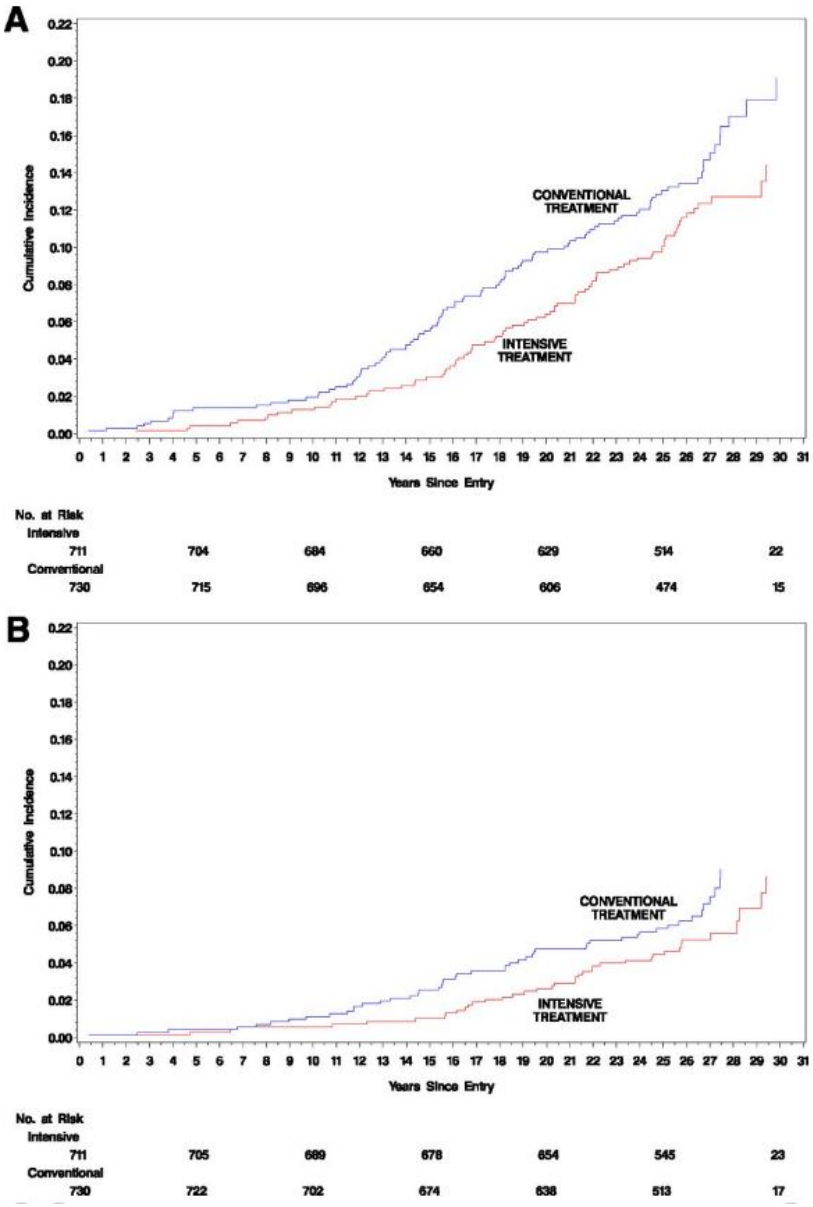
Event	Initial events n Manuscript	Initial events n DSIC	Diff.	Secondary events n Manuscript	Secondary events n DSIC	Diff.
6. Revascularization	22	22	0	49	49	0
7. CHF	3	3	0	10	10	0
MACE	49	49	0	15	15	0
1. Nonfatal acute MI	28	28	0	7	7	0
2. Nonfatal cerebrovascular event	11	11	0	2	2	0
3. Death from CVD	10	10	0	6	6	0

Table C: Variables used to replicated Figure 1

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

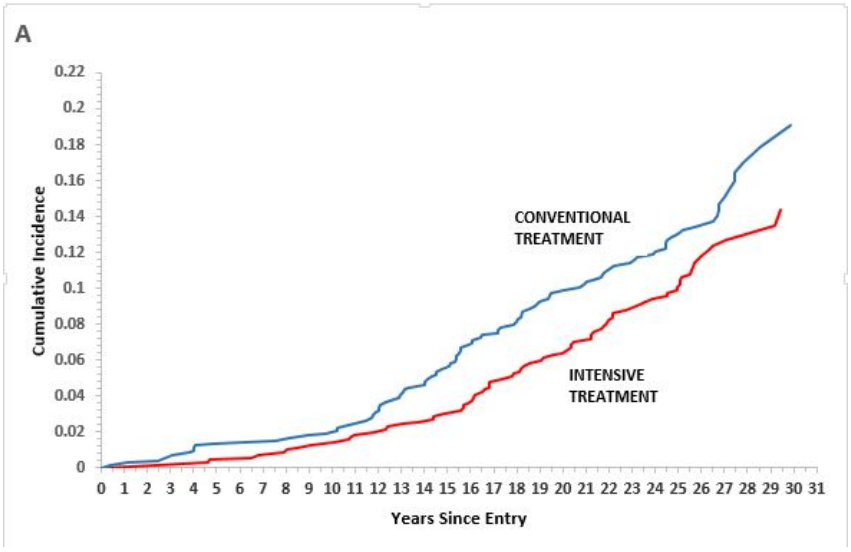
Figure A: Comparison of values computed in integrity check to reference article Figure 1 values

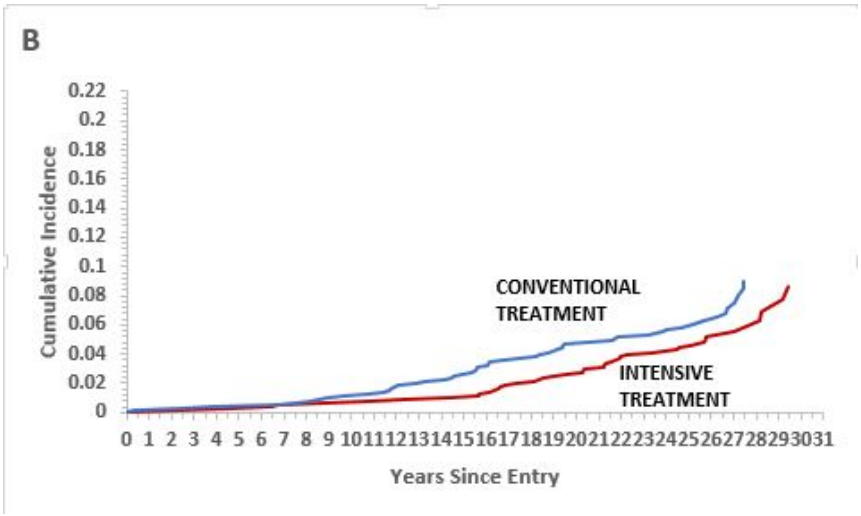
Manuscript:



Cumulative incidence of cardiovascular outcomes in the conventional treatment and intensive treatment groups during up to 30 years of DCCT/EDIC treatment and follow-up. A: The first of any of the predefined CVD outcomes. The risk reduction with intensive therapy was 30% (95% CI 7, 48; P = 0.016). B: The first occurrence of MACE. The risk reduction with intensive therapy was 32% (95% CI -3, 56; P = 0.07).

DSIC:


A
Cumulative Incidence
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
CONVENTIONAL TREATMENT
INTENSIVE TREATMENT
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Years Since Entry


B
Cumulative Incidence
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
CONVENTIONAL TREATMENT
INTENSIVE TREATMENT
0 1 2 3 4 5 6 7 8 9 10111213141516171819202122232425262728293031
Years Since Entry

A:

No. at Risk	Year 0 Manuscript	Year 0 DSIC	Diff.	Year 5 Manuscript	Year 5 DSIC	Diff.	Year 10 Manuscript	Year 10 DSIC	Diff.
Intensive	711	711	0	704	704	0	684	684	0
Conventional	730	730	0	715	715	0	696	696	0

No. at Risk	Year 15 Manuscript	Year 15 DSIC	Diff.	Year 20 Manuscript	Year 20 DSIC	Diff.	Year 25 Manuscript	Year 25 DSIC	Diff.
Intensive	660	665	5	629	629	0	514	514	0
Conventional	654	654	0	606	606	0	474	474	0

No. at Risk	Year 30 Manuscript	Year 30 DSIC	Diff.
Intensive	22	22	0
Conventional	15	15	0

B:

No. at Risk	Year 0 Manuscript	Year 0 DSIC	Diff.	Year 5 Manuscript	Year 5 DSIC	Diff.	Year 10 Manuscript	Year 10 DSIC	Diff.
Intensive	711	711	0	705	705	0	689	689	0
Conventional	730	730	0	722	722	0	702	702	0

No. at Risk	Year 15 Manuscript	Year 15 DSIC	Diff.	Year 20 Manuscript	Year 20 DSIC	Diff.	Year 25 Manuscript	Year 25 DSIC	Diff.
Intensive	678	678	0	654	654	0	545	545	0
Conventional	674	674	0	638	638	0	513	513	0

No. at Risk	Year 30 Manuscript	Year 30 DSIC	Diff.
Intensive	23	23	0
Conventional	17	17	0

Attachment A: SAS Code

```
*** EDIC CVD Data DSIC;
*** 30 yr outcome data;
*** Programmer: Allyson Mateja;
*** Date: 12/20/2016;

title 'DCCT/EDIC CVD 30 yr Data DSIC';
title2 ' ';

proc format;
       value cvdf 1 = 'Nonfatal acute MI'
                  2 = 'Nonfatal cerbrovascular event'
                  3 = 'Death from CVD'
                  4 = 'Silent MI'
                  5 = 'Confirmed angina'
                  6 = 'Revascularization'
                  7 = 'CHF';

       value initialf 0 = 'Secondary event'
                      1 = 'Initial event';

libname cvddata '/prj/niddk/ims_analysis/DCCT_EDIC/private_orig_data/CVD_Datal/';

data cvd_all_subjects;
       set cvddata.nih_cvd30yr_1441;

data cvd_events;
       set cvddata.nih_cvd30yr_ev_366;

proc contents data = cvd_all_subjects;
proc contents data = cvd_events;

proc freq data = cvd_all_subjects;
       tables group;

proc sort data = cvd_events;
       by mask_pat;

proc sort data = cvd_all_subjects;
       by mask_pat;

data cvd_events;
       merge cvd_all_subjects (in=val1 keep=mask_pat group hard carvtime1 hardtime1)
             cvd_events       (in=val2);
       by mask_pat;
       if val1 and val2 then output;

proc sort data = cvd_events nodupkey out = subjects;
       by mask_pat cvd_event;
```

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

data mask_pat_only;
       merge mask_pat_only    (in=val1)
             cvd_all_subjects (in=val2 keep=mask_pat group hard);
       by mask_pat;
       if val1 then any_cvd_event = 1;
       else any_cvd_event = 0;
       if val2 then output;

proc sort data = mask_pat_only;
       by group;

proc freq data = mask_pat_only;
       tables any_cvd_event;
       by group;
       title3 'Table 1 - Patients (n) Any CVD Event';

proc freq data = mask_pat_only;
       tables any_cvd_event*hard /list;
       by group;
       title3 'Table 1 - MACE Patients (n) Any CVD Event';

data acute_mi nonfatal_cvd_event death_cvd silent_mi angina revasc chf;
       set subjects;
       if cvd_event=1 then output acute_mi;
       if cvd_event=2 then output nonfatal_cvd_event;
       if cvd_event=3 then output death_cvd;
       if cvd_event=4 then output silent_mi;
       if cvd_event=5 then output angina;
       if cvd_event=6 then output revasc;
       if cvd_event=7 then output chf;

data subjects;
       merge acute_mi           (in=val1 keep=mask_pat)
             nonfatal_cvd_event (in=val2 keep=mask_pat)
             death_cvd          (in=val3 keep=mask_pat)
             silent_mi          (in=val4 keep=mask_pat)
             angina             (in=val5 keep=mask_pat)
             revasc             (in=val6 keep=mask_pat)
             chf                (in=val7 keep=mask_pat)
             cvd_all_subjects   (in=val8 keep=mask_pat group hard);
       by mask_pat;
       if val1 then mi_event = 1;
       else mi_event = 0;
       if val2 then nonfatal_cvd = 1;
       else nonfatal_cvd = 0;
       if val3 then death_cvd = 1;
       else death_cvd = 0;
```

```
        if val4 then silent_mi = 1;
        else silent_mi = 0;
        if val5 then angina = 1;
        else angina = 0;
        if val6 then revascularization = 1;
        else revascularization = 0;
        if val7 then chf = 1;
        else chf = 0;
        if val8 then output;

data mace_events;
        set cvd_events;
        if hard=1 and cvd_event in (1,2,3);

proc sort data = mace_events;
        by mask_pat cvd_event;

data mace_events;
        set mace_events;
        by mask_pat;
        mace_init_event = 0;
        if first.mask_pat and mask_pat ne 1214 then mace_init_event = 1;
        if init_event = 1 and mace_init_event=0 then mace_init_event=1;

proc sort data = mace_events;
        by mask_pat cvd_event;

proc sort data = cvd_events;
        by mask_pat cvd_event;

data cvd_events;
        merge cvd_events  (in=val1)
              mace_events (keep=mask_pat cvd_event mace_init_event);
        by mask_pat cvd_event;
        if val1 then output;

proc sort data = subjects;
        by group;

proc freq data = subjects;
        tables mi_event nonfatal_cvd death_cvd silent_mi angina revascularization chf;
        by group;
        title3 'Table 1 - Patients (n) by Treatment Group';

proc freq data = subjects;
        tables mi_event*hard nonfatal_cvd*hard death_cvd*hard /list;
        by group;
        title3 'Table 1 - MACE Patients (n) by Treatment Group';

proc sort data = cvd_events;
        by group;
```

```
proc freq data = cvd_events;
       tables cvd_event;
       format cvd_event cvdf.;
       by group;
       title3 'Table 1 - Events by Treatment Group';

proc freq data = cvd_events;
       tables cvd_event;
       format cvd_event cvdf.;
       by group;
       where hard=1 and cvd_event in (1,2,3);
       title3 'Table 1 - MACE Events by Treatment Group';

proc freq data = cvd_events;
       tables init_event cvd_event*init_event /list;
       format cvd_event cvdf. init_event initialf.;
       by group;
       title3 'Table 1 - Events by Treatment Group';

proc freq data = cvd_events;
       tables mace_init_event cvd_event*mace_init_event /list;
       format cvd_event cvdf. mace_init_event initialf.;
       by group;
       where hard=1 and cvd_event in (1,2,3);
       title3 'Table 1 - MACE Events by Treatment Group';

ods graphics on;

ods csv file = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part1.csv';
PROC LIFETEST DATA=cvd_all_subjects plots=survival(atrisk=0 to 31 by 5 nocensor failure);
TIME  carvtime1 * carv (0);
strata group;

RUN;
ods csv close;

proc import datafile = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part1.xls'
       dbms = xls
       out = outcomes_experimental;
       sheet = 'Experimental';
       getnames = yes;
run;

proc import datafile = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part1.xls'
       dbms = xls
       out = outcomes_standard;
       sheet = 'Standard';
       getnames = yes;
run;
```

```
proc sort data = outcomes_experimental nodupkey;
       by years failure_intensive;

proc sort data = outcomes_standard nodupkey;
       by years failure_standard;

data outcomes_both;
       merge outcomes_experimental (in=val1)
             outcomes_standard     (in=val2);
       by years;
       if val1 or val2 and (failure_intensive ne . or failure_standard ne .) then output;

ods csv file = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part1_v2.csv';
       proc print data = outcomes_both noobs;
run;
ods csv close;


ods csv file = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part2.csv';
PROC LIFETEST DATA=cvd_all_subjects plots=survival(atrisk=0 to 31 by 5 nocensor failure);
TIME  hardtime1 * hard (0);
strata group;
run;
ods csv close;


proc import datafile = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part2.xls'
       dbms = xls
       out = outcomes_experimental
       replace;
       sheet = 'Experimental';
       getnames = yes;
run;

proc import datafile = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part2.xls'
       dbms = xls
       out = outcomes_standard
       replace;
       sheet = 'Standard';
       getnames = yes;
run;

proc sort data = outcomes_experimental nodupkey;
       by years failure_intensive;

proc sort data = outcomes_standard nodupkey;
       by years failure_standard;

data outcomes_both;
       merge outcomes_experimental (in=val1)
             outcomes_standard     (in=val2);
```

```
        by years;
        if val1 or val2 and (failure_intensive ne . or failure_standard ne .) then output;

ods csv file = '/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/cvd_dsic_part2_v2.csv';
        proc print data = outcomes_both noobs;
run;
ods csv close;
```