# Dataset Integrity Check for DCCT/EDIC EDRET18 Data File

**Prepared by Allyson Mateja**
3901 Calverton Blvd, Suite 200 Calverton MD 20705
**December 10, 2015**

# 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 file edret18.sas7bdat.

# 4 Statistical Methods

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

# 5 Results

For Table 1 in the publication [1], Table 1. Clinical Characteristics of the former DCCT INT and CONV participants at DCCT baseline, DCCT closeout, and EDIC years 15-18, Table A lists the variables that can be used in the replication. Table C compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are similar to the results in publication [1].

For Table 2 in the publication [1], Incidence of Further three or more-step progression of retinopathy and new PDR between the end of the DCCT and after 18 years of the EDIC study overall and stratified by the level of retinopathy at the end of DCCT, Table D compares the results calculated from the archived data file to the results published in Table 2. The results of the replication are similar to the results in publication [1].

For Table 3 in the publication [1], Prevalence of Various retinopathy complications in the former DCCT INT and CONV at DCCT closeout, EDIC year 10, and EDIC years 15-18 among 1,214 patients evaluated for retinopathy or CSME during EDIC years 15-18, Table E lists the variables that can be used in the replication. Table G compares the results calculated from the archived data file to the results published in Table 3. The results of the replication are similar to the results in publication [1].

# 6 Conclusions

The NIDDK repository is confident that the EDIC data files to be distributed is a true copy to the manuscript data when available.

# 7 References

[1] Lachin JM, et al.; DCCT/EDIC Research Group. Effect of Intensive Diabetes Therapy on the Progression of Diabetic Retinopathy in Patients With Type 1 Diabetes: 18 Years of Follow-up in the DCCT/EDIC. Diabetes 2015;64:631-642.

**Table A:** Variables used to replicate Table 1 Clinical Characteristics of the former DCCT INT and CONV participants at DCCT baseline, DCCT closeout, and EDIC years 15-18

| Table Variable | Variable Used in Replication from Dataset |
|---|---|
| | |
| Age (years) | age |
| Female (%) | sex |
| Diabetes Duration | fattdury |
| DCCT primary cohort | primary |
| Hypertension | fht |
| Hyperlipidemia | fhlip |
| Current cigarette smoking | fmsokes |
| Pump or MDI | fmdi |
| Glucose Monitoring More than 4 Times a Day | fgluc4 |
| Use of ACE inhibitor or ARB | acearb |
| BMI | fbmi |
| Obese | fobese |
| Systolic BP | fsbp |
| Diastolic BP | fdbp |
| Mean arterial pressure | fbpm |
| HbA1c (%) | fhba1c (DCCT baseline),<br>dcct_hba (DCCT end),<br>fedic_hba (EDIC years 15-18) |
| HbA1c (mmol/mol) | ((hba1c - 2.152) * 10.931) |
| Total cholesterol | fchl |
| HDL cholesterol | fhdl |
| LDL cholesterol | fldl |
| Triglycerides | ftrg |
| No Retinopathy | ret, etdrscat3 (=1) |
| MA Only | fetdrspat (= 2 or 3), etdrscat3 (=2) |
| Mild NPDR | fetdrspat (= 4 or 5), etdrscat3 (=3) |
| Moderate NPDR | fetdrspat (= 6, 7, 8, or 9), etdrscat3 (=4) |
| SNPDR | snpdr, etdrscat3 (>=5) |
| Sustained AER > 30 | fsaer30 |
| AER > 300 | faer300 |
| Sustained eGFR < 60 | scgfr60 |

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

DCCT Baseline

| | INT Manuscript | INT DSIC | Difference | CONV Manuscript | CONV DSIC | Difference |
|---|---|---|---|---|---|---|
| n | 711 | 711 | 0 | 730 | 730 | 0 |
| Age(years) | 27.2 (7.1) | 27.2 (7.1) | 0 | 26.7 (7.1) | 26.7 (7.1) | 0 |
| Female | 48.5 | 48.5 | 0 | 45.9 | 45.9 | 0 |
| Diabetes Duration | 5.8 (4.2) | 5.8 (4.2) | 0 | 5.5 (4.1) | 5.5 (4.1) | 0 |
| DCCT primary cohort | 49 | 48.9 | 0.1 | 51.8 | 51.8 | 0 |
| Hypertension | 3.1 | 3.1 | 0 | 2.1 | 2.1 | 0 |
| Hyperlipidemia | 22.8 | 22.8 | 0 | 23.4 | 23.4 | 0 |
| Current cigarette smoking | 18.6 | 18.6 | 0 | 18.4 | 18.4 | 0 |
| Pump or MDI | 0 | 0 | 0 | 0 | 0 | 0 |
| Glucose Monitoring More than 4 Times a Day | 0 | 0 | 0 | 0 | 0 | 0 |
| Use of ACE inhibitor or ARB | 0 | 0 | 0 | 0 | 0 | 0 |
| BMI | 23.4 (2.7) | 23.4 (2.7) | 0 | 23.5 (2.9) | 23.5 (2.9) | 0 |
| Obese | 1.3 | 1.3 | 0 | 1.9 | 1.9 | 0 |
| Systolic BP | 114.5 (11.3) | 114.5 (11.3) | 0 | 114.6 (11.4) | 114.6 (11.4) | 0 |
| Diastolic BP | 73.1 (8.2) | 73.1 (8.2) | 0 | 72.9 (8.7) | 72.9 (8.7) | 0 |
| Mean arterial pressure | 86.9 (8.2) | 86.9 (8.2) | 0 | 86.8 (8.6) | 86.8 (8.6) | 0 |
| HbA1c (%) | 9.1 (1.6) | 9.1 (1.6) | 0 | 9.1 (1.6) | 9.1 (1.6) | 0 |
| HbA1c (mmol/mol) | 76 (17.5) | 75.8 (17.4) | 0.2 (0.1) | 76 (17.5) | 75.5 (17.9) | 0.5 (0.4) |
| Total cholesterol | 177.1 (32.8) | 177.1 (32.8) | 0 | 175.7 (33.6) | 175.7 (33.6) | 0 |
| HDL cholesterol | 50.8 (12.3) | 50.8 (12.3) | 0 | 50.3 (12.3) | 50.3 (12.3) | 0 |
| LDL cholesterol | 110.3 (28.7) | 110.3 (28.7) | 0 | 109.1 (29.4) | 109.2 (29.4) | 0.1 (0) |
| Triglycerides | 80.8 (43.3) | 80.8 (43.3) | 0 | 81.8 (51.3) | 81.8 (51.3) | 0 |
| No Retinopathy | 48.9 | 48.9 | 0 | 51.8 | 51.8 | 0 |
| MA Only | 35.1 | 35 | 0.1 | 27.8 | 27.8 | 0 |
| Mild NPDR | 11.6 | 11.5 | 0.1 | 15.2 | 15.2 | 0 |
| Moderate NPDR | 4.5 | 4.5 | 0 | 5.1 | 5.1 | 0 |
| SNPDR | 0 | 0 | 0 | 0.1 | 0.1 | 0 |
| Sustained AER > 30 | 5.2 | 5.2 | 0 | 4.3 | 4.2 | 0.1 |
| AER > 300 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sustained eGFR < 60 | 0 | 0 | 0 | 0 | 0 | 0 |

End of DCCT

| | INT Manucript | INT DSIC | Difference | CONV Manuscript | CONV DSIC | Difference |
|---|---|---|---|---|---|---|
| n | 701 | 701 | 0 | 722 | 722 | 0 |
| Age(years) | 33.6 (7.0) | 33.6 (7.0) | 0 | 33.0 (7.0) | 33.0 (7.0) | 0 |
| Female | 48.9 | 48.9 | 0 | 46 | 46 | 0 |
| Diabetes Duration | 12.3 (4.9) | 12.3 (4.9) | 0 | 11.9 (4.8) | 11.9 (4.8) | 0 |
| DCCT primary cohort | 49.2 | 49.2 | 0 | 51.7 | 51.7 | 0 |
| Hypertension | 4.4 | 4.4 | 0 | 3.9 | 3.9 | 0 |
| Hyperlipidemia | 25.8 | 25.8 | 0 | 29.9 | 29.9 | 0 |
| Current cigarette smoking | 20.3 | 20.3 | 0 | 19.8 | 19.8 | 0 |
| Pump or MDI | 97.2 | 97.1 | 0.1 | 5.1 | 5.1 | 0 |
| Glucose Monitoring More than 4 Times a Day | 52.6 | 52.6 | 0 | 3.7 | 3.7 | 0 |
| Use of ACE inhibitor or ARB | N/A | N/A | N/A | N/A | N/A | N/A |
| BMI | 26.5 (4.2) | 26.5 (4.2) | 0 | 25.0 (3.1) | 25.0 (3.1) | 0 |
| Obese | 18.5 | 18.5 | 0 | 5.7 | 5.7 | 0 |
| Systolic BP | 116.3 (11.7) | 116.3 (11.7) | 0 | 115.3 (12.0) | 115.3 (12.0) | 0 |
| Diastolic BP | 74.4 (8.8) | 74.4 (8.8) | 0 | 74.2 (8.8) | 74.2 (8.8) | 0 |
| Mean arterial pressure | 88.3 (8.9) | 88.3 (8.9) | 0 | 87.9 (8.9) | 87.9 (8.9) | 0 |
| HbA1c (%) | 7.2 (0.9) | 7.2 (0.9) | 0 | 9.1(1.3) | 9.1 (1.3) | 0 |
| HbA1c (mmol/mol) | 55 (9.8) | 55.7 (10.1) | 0.7 (0.3) | 76 (14.2) | 75.9 (13.8) | 0.1 (0.4) |
| Total cholesterol | 178.9 (31.3) | 178.9 (31.4) | 0 (0.1) | 183.7 (36.9) | 183.7 (36.9) | 0 |
| HDL cholesterol | 50.8 (12.8) | 50.8 (12.8) | 0 | 51.6 (12.9) | 51.6 (12.9) | 0 |
| LDL cholesterol | 111.7 (27.3) | 111.7 (27.3) | 0 | 114.6 (31.5) | 114.6 (31.5) | 0 |
| Triglycerides | 81.9 (51.5) | 81.9 (51.5) | 0 | 88.3 (54.5) | 88.3 (54.5) | 0 |
| No Retinopathy | 28.3 | 28.3 | 0 | 17.2 | 17.2 | 0 |
| MA Only | 39.7 | 39.7 | 0 | 32.1 | 32.1 | 0 |
| Mild NPDR | 21.3 | 21.3 | 0 | 28.5 | 28.5 | 0 |
| Moderate NPDR | 8.3 | 8.1 | 0.2 | 14.4 | 14.4 | 0 |
| SNPDR | 2.6 | 2.6 | 0 | 7.8 | 7.8 | 0 |
| Sustained AER > 30 | 7.6 | 7.6 | 0 | 14.5 | 14.5 | 0 |
| AER > 300 | 1.4 | 1.4 | 0 | 3.2 | 3.2 | 0 |
| Sustained eGFR < 60 | 0.1 | 0.1 | 0 | 0.4 | 0.4 | 0 |

EDIC Years 15-18

| | INT paper | INT DSIC | Difference | CONV paper | CONV DSIC | Difference |
|---|---|---|---|---|---|---|
| n | 606 | 606 | 0 | 608 | 608 | 0 |
| Age(years) | 50.9 (7.2) | 50.9 (7.2) | 0 | 49.9 (7.0) | 49.9 (7.0) | 0 |
| Female | 48.8 | 48.8 | 0 | 46.9 | 46.9 | 0 |
| Diabetes Duration | 29.3 (5.3) | 29.4 (5.3) | 0.1 (0) | 28.7 (5.4) | 28.8 (5.4) | 0.1 (0) |
| DCCT primary cohort | 48.4 | 48.3 | 0.1 | 50.8 | 50.8 | 0 |
| Hypertension | 62.4 | 62.5 | 0.1 | 66 | 66.1 | 0.1 |
| Hyperlipidemia | 64.5 | 64.5 | 0 | 66.8 | 66.8 | 0 |
| Current cigarette smoking | 12.2 | 12 | 0.2 | 12.2 | 12.2 | 0 |
| Pump or MDI | 98.2 | 98.2 | 0 | 96.1 | 96.1 | 0 |
| Glucose Monitoring More than 4 Times a Day | 66.8 | 66.9 | 0.1 | 70.2 | 70.4 | 0.2 |
| Use of ACE inhibitor or ARB | 53 | 53.1 | 0.1 | 57.6 | 57.9 | 0.3 |
| BMI | 28.9 (5.6) | 28.9 (5.6) | 0 | 28.2 (5.0) | 28.2 (5.0) | 0 |
| Obese | 35.6 | 35.8 | 0.2 | 31.4 | 31.4 | 0 |
| Systolic BP | 121.1 (14.5) | 121.2 (14.6) | 0.1 (0.1) | 120.4 (14.7) | 120.4 (14.6) | 0 (0.1) |
| Diastolic BP | 71.7 (9.0) | 71.6 (9.0) | 0.1 (0) | 71.3 (8.8) | 71.3 (8.7) | 0 (0.1) |
| Mean arterial pressure | 88.1 (9.5) | 88.2 (9.5) | 0.1 (0) | 87.7 (9.4) | 87.7 (9.4) | 0 |
| HbA1c (%) | 8.0 (1.1) | 8.0 (1.1) | 0 | 8.0 (1.0) | 8.0 (1.0) | 0 |
| HbA1c (mmol/mol) | 64 (12.0) | 63.5 (11.7) | 0.5 (0.3) | 64 (10.9) | 63.8 (11) | 0.2 (0.1) |
| Total cholesterol | 175.4 (36.2) | 175.4 (36.1) | 0 (0.1) | 172.5 (38.5) | 172.5 (38.5) | 0 |
| HDL cholesterol | 61.3 (19.4) | 61.3 (19.4) | 0 | 61.6 (18.3) | 61.6 (18.1) | 0 (0.2) |
| LDL cholesterol | 97.3 (29.5) | 97.3 (29.4) | 0 (0.1) | 94.4 (30.5) | 94.4 (30.5) | 0 |
| Triglycerides | 84.4 (54.9) | 84.3 (54.7) | 0.1 (0.2) | 83.4 (76.7) | 83.3 (76.5) | 0.1 (0.2) |
| No Retinopathy | 10.8 | 10.8 | 0 | 4.8 | 4.8 | 0 |
| MA Only | 36.9 | 36.6 | 0.3 | 26.6 | 26.2 | 0.4 |
| Mild NPDR | 20.2 | 20.2 | 0 | 18.1 | 18.0 | 0.1 |
| Moderate NPDR | 16.5 | 16.4 | 0.1 | 19.7 | 19.2 | 0.5 |
| SNPDR | 15.5 | 15.9 | 0.4 | 31.2 | 31.3 | 0.1 |
| Sustained AER > 30 | 13.5 | 13.5 | 0 | 20.6 | 20.6 | 0 |
| AER > 300 | 4 | 4.1 | 0.1 | 7.4 | 7.4 | 0 |
| Sustained eGFR < 60 | 3.9 | 3.9 | 0 | 5.4 | 5.4 | 0 |

**Table C:** Variables used to replicate Table 2 Incidence of Further three or more-step progression of retinopathy and new PDR between the end of the DCCT and after 18 years of the EDIC study overall and stratified by the level of retinopathy at the end of DCCT

| Table Variable | Variable Used in Replication from Dataset |
|---|---|
| | |
| Stratum 1: No retinopathy | dcct10 |
| Stratum 2: MA Only | dcct20 |
| Stratum 3: Mild NPDR | dcct30 |
| Stratum 4: Moderate or severe NPDR | dcct40 |
| Further >= 3-step progression | anystp3f |
| PDR | anypdr |

**Table D:** Comparison of values computed in integrity check to reference article Table 2 values

| | n at risk (Further 3 Step Progression) | DSIC | Difference | n with event - Further 3 Step Progression (%) | DSIC | Difference |
|---|---|---|---|---|---|---|
| All Levels | 1358 | 1357 | 1 | | | |
| Int | 684 | 684 | 0 | 267(39) | 267 (39) | 0 |
| Conv | 674 | 673 | 1 | 380 (56.4) | 380 (56.5) | 0 (0.1) |
| No Retinopathy | | | | | | |
| Int | 194 | 194 | 0 | 100 (51.6) | 100 (51.6) | 0 |
| Conv | 123 | 123 | 0 | 74 (60.2) | 74 (60.2) | 0 |
| MA Only | | | | | | |
| Int | 275 | 275 | 0 | 88 (32) | 85 (32) | 0 |
| Conv | 220 | 220 | 0 | 112 (50.9) | 112 (50.9) | 0 |
| Mild NPDR | | | | | | |
| Int | 149 | 149 | 0 | 44 (29.5) | 44 (29.5) | 0 |
| Conv | 200 | 200 | 0 | 101 (50.5) | 101 (50.5) | 0 |
| Moderate NPDR | | | | | | |
| Int | 65 | 65 | 0 | 35 (53.9) | 35 (53.8) | 0 (0.1) |
| Conv | 126 | 126 | 0 | 93 (73.8) | 93 (73.8) | 0 |

| | n at risk (PDR) | DSIC | Difference | n with event – PDR (%) | DSIC | Difference |
|---|---|---|---|---|---|---|
| All Levels | 1318 | 1317 | 1 | | | |
| Int | 668 | 668 | 0 | 86 (12.9) | 86 (12.9) | 0 |
| Conv | 650 | 649 | 1 | 172 (26.5) | 172 (26.5) | 0 |
| No Retinopathy | | | | | | |
| Int | 194 | 194 | 0 | 8 (4.1) | 8 (4.1) | 0 |
| Conv | 122 | 122 | 0 | 5 (4.1) | 5 (4.1) | 0 |
| MA Only | | | | | | |
| Int | 274 | 274 | 0 | 23 (8.4) | 23 (8.4) | 0 |
| Conv | 220 | 220 | 0 | 32 (14.6) | 32 (14.6) | 0 |
| Mild NPDR | | | | | | |
| Int | 149 | 149 | 0 | 31 (20.8) | 31 (20.8) | 0 |
| Conv | 199 | 199 | 0 | 64 (32.2) | 64 (32.2) | 0 |
| Moderate NPDR | | | | | | |
| Int | 50 | 50 | 0 | 24 (48) | 24 (48) | 0 |
| Conv | 104 | 104 | 0 | 71 (68.3) | 71 (68.3) | 0 |

**Table E:** Variables used to replicate Table 3 Prevalence of Various retinopathy complications in the former DCCT INT and CONV at DCCT closeout, EDIC year 10, and EDIC years 15-18 among 1,214 patients evaluated for retinopathy or CSME during EDIC years 15-18

| Table Variable | Variable Used in Replication from Dataset |
|---|---|
| | |
| 3 step progression from DCCT baseline | step3 |
| SNPDR | snpdr |
| PDR | pdr |
| csme | csme |
| Photocoagulation Therapy | anyfocal3, anyscat3, anyavegf |

**Table G:** Comparison of values computed in integrity check to reference article Table 3 values

| | DCCT Closeout | | | | | | EDIC year 10 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | INT | INT DSIC | Difference | CONV | CONV DSIC | Difference | INT | INT DSIC | Difference | CONV | CONV DSIC | Difference |
| n | 606 | 606 | 0 | 608 | 608 | 0 | 559 | 559 | 0 | 574 | 574 | 0 |
| 3 step progression from DCCT baseline | 9.2 | 9.4 | 0.2 | 31.5 | 31.6 | 0.1 | 34.3 | 34.5 | 0.2 | 60.6 | 60.9 | 0.3 |
| SNPDR | 2.2 | 2.3 | 0.1 | 8.3 | 8.4 | 0.1 | 8.8 | 8.8 | 0 | 25.8 | 26.2 | 0.4 |
| PDR | 2 | 2.1 | 0.1 | 7 | 6.9 | 0.1 | 8.6 | 8.7 | 0.1 | 25.4 | 25.8 | 0.4 |
| | | | | | | | | | | | | |
| n | 581 | 581 | 0 | 562 | 562 | 0 | 533 | 533 | 0 | 521 | 521 | 0 |
| CSME | 3.8 | 3.8 | 0 | 6.8 | 6.9 | 0.1 | 9.4 | 9.4 | 0 | 20.4 | 20.3 | 0.1 |
| | | | | | | | | | | | | |
| n | 606 | 606 | 0 | 608 | 608 | 0 | 559 | 559 | 0 | 574 | 574 | 0 |
| Photocoagulation Therapy | 3.3 | 3.3 | 0 | 7.7 | 7.7 | 0 | 9.3 | 9.3 | 0 | 25.4 | 25.4 | 0 |

| | EDIC years 15-18 | | | | | |
|---|---|---|---|---|---|---|
| | INT | INT DSIC | Difference | CONV | CONV DSIC | Difference |
| n | 606 | 606 | 0 | 608 | 608 | 0 |
| 3 step progression from DCCT Baseline | 41.1 | 40.6 | 0.5 | 58.7 | 58.3 | 0.4 |
| SNPDR | 15.9 | 15.9 | 0 | 31.5 | 31.3 | 0.2 |
| PDR | 15.7 | 15.8 | 0.1 | 31.5 | 31.3 | 0.2 |
| | | | | | | |
| n | 581 | 581 | 0 | 562 | 562 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | EDIC years 15-18 | | | | | |
| CSME | 17 | 17 | 0 | 26 | 26 | 0 |
| | | | | | | |
| n | 606 | 606 | 0 | 608 | 608 | 0 |
| Photocoagulation Therapy | 17.2 | 17.8 | 0.6 | 30.9 | 31.7 | 0.8 |

# Attachment A: SAS Code

```
title 'NIDDK DCCT EDIC edret18 Dataset';
/*options symbolgen mprint;*/

/* Created by: Allyson Mateja

Date: July 2015

Modeled after CMRI DSIC by Michael Spriggs

Prepared for DCCT-EDIC Dataset Integrity Check*/

%global caser;
%global timeline;

%macro freqdata1 (order=, invar=, level=);

data data0 data1;
       set _null_;

proc freq data=table1 noprint ;
       tables &invar /out=data0 sparse;
       by &caser;
       where &timeline;
       format _all_;
       run;

data data1;
       set data0;
       length LEVEL $100;
       LEVEL=strip(&invar);

data data1 (keep = LEVEL &caser name CHARALL ORDERER);
       set data1;
       length name $100 CHARALL $100;
       name = "&invar";
       PCT_DISP = round(PERCENT, .1);
       CHARALL = /*compress(put(COUNT, 8.))||'('||*/compress(put(PCT_DISP, 8.1))/*||')'*/;
       ORDERER = &order;
       if level in &level then output data1;

data accumfreq1;
       set accumfreq1 data1;

%mend freqdata1;

%macro freqdata2 (order=, invar1=, invar2=,  level=);

data data0 data1;
```

```
        set _null_;

proc freq data=table1 noprint;
        tables &invar1 /out=data0;
        by &caser;
        where &invar2;
        format _all_;
        run;

data data1;
        set data0;
        length LEVEL $100;
        LEVEL=strip(&invar1);

data data1 (keep=LEVEL &caser name name2 CHARALL ORDERER);
        set data1;
        length name $100 CHARALL $100;
        length name2 $100 CHARALL $100;
        name = "&invar1";
        name2 = "&invar2";
        PCT_DISP = round (PERCENT, .1);
        CHARALL = compress(put(COUNT, 8.))||' ('||compress(put(PCT_DISP, 8.1))||')';
        ORDERER = &order;
        if level in &level then output data1;

data accumfreq2;
        set accumfreq2 data1;

%mend freqdata2;

%macro freqdata3 (order=, invar=);

data data0 data1;
        set _null_;

proc freq data=table1 noprint ;
        tables &caser /out=data0 sparse;
        where &invar;
        format _all_;
        run;

data data1;
        set data0;

data data1 (keep = &caser name CHARALL ORDERER);
        set data1;
        length name $100 CHARALL $100;
        name = "&invar";
        PCT_DISP = round(PERCENT, .1);
        CHARALL = compress(put(COUNT, 8.))/*||'('||compress(put(PCT_DISP, 8.1))||')'*/;
        ORDERER = &order;
```

```
data accumfreq3;
	set accumfreq3 data1;

%mend freqdata3;

%macro meandata1 (order=, invar=, roundvar=, digit=);
proc means data=table1 mean stddev noprint;
	where &timeline;
	var &invar;
	class &caser;
	output out=data1
			mean=mean
			stddev=stddev;
	run;

data data1 (drop=_TYPE_ _FREQ_ mean stddev);
	set data1;
	length name CHARALL $100;
	name="&invar";
	mean = round(mean, &roundvar);
	stddev=round(stddev, &roundvar);
	CHARALL = compress(put(mean, 8.&digit))||'(' ||compress(put(stddev, 8.&digit))||')';
	ORDERER = &order;

data accummean1;
	set accummean1 data1;

%mend meandata1;

%macro datachunk();

%meandata1(order=1, invar = age, roundvar = .1, digit =1);
%freqdata1(order=2, invar = sex, level = ('F'));
%meandata1(order=3, invar=fattdury, roundvar = .1, digit = 1);
%freqdata1(order=5, invar=primary, level = ('1'));
%freqdata1(order=6, invar=fht, level = ('1'));
%freqdata1(order=8, invar=fhlip, level = ('1'));
%freqdata1(order=10, invar=fsmokes, level = ('1'));
%freqdata1(order=13, invar=fmdi, level = ('1'));
%freqdata1(order=14, invar=fgluc4, level = ('1'));
%freqdata1(order=16, invar=acearb, level = ('1'));
%meandata1(order=17, invar=fbmi, roundvar = .1, digit = 1);
%freqdata1(order=20, invar=fobese, level = ('1'));
%meandata1(order=21, invar=fsbp, roundvar=.1, digit=1);
%meandata1(order=23, invar=fdbp, roundvar=.1, digit=1);
%meandata1(order=25, invar=fbpm, roundvar=.1, digit=1);
%meandata1(order=26, invar=fhba1c, roundvar=.1, digit=1);
%meandata1(order=27, invar=hbammol, roundvar=.1, digit=1);
%meandata1(order=28, invar=dcct_hba, roundvar=.1, digit=1);
%meandata1(order=28.1, invar=hbammol_dcct, roundvar=.1, digit=1);
```

```
%meandata1(order=30, invar=fedic_hba, roundvar=.1, digit=1);
%meandata1(order=30.1, invar=hbammol_edic, roundvar=.1, digit=1);
%meandata1(order=35, invar=fchl, roundvar=.1, digit=1);
%meandata1(order=37, invar=fhdl, roundvar=.1, digit=1);
%meandata1(order=39, invar=fldl, roundvar=.1, digit=1);
%meandata1(order=41, invar=ftrg, roundvar=.1, digit=1);
%freqdata1(order=42, invar=ret, level = ('0'));
%freqdata1(order=43, invar=ma_only, level = ('1'));
%freqdata1(order=44, invar=mildnpdr, level = ('1'));
%freqdata1(order=45, invar=moderatenpdr, level = ('1'));
%freqdata1(order=46, invar=snpdr, level = ('1'));
%freqdata1(order=46.1, invar=etdrscat3, level = ('1'));
%freqdata1(order=46.2, invar=etdrscat3, level = ('2'));
%freqdata1(order=46.3, invar=etdrscat3, level = ('3'));
%freqdata1(order=46.4, invar=etdrscat3, level = ('4'));
%freqdata1(order=46.5, invar=etdrscat3, level = ('5'));
%freqdata1(order=48, invar=fsaer30, level = ('1'));
%freqdata1(order=49, invar=faer300, level = ('1'));
%freqdata1(order=51, invar=scgfr60, level = ('1'));



%mend datachunk;

%macro table2_datachunk();

%freqdata3(order = 1, invar = dcct10);
%freqdata2(order = 2, invar1 = dcct10, invar2 = ANYSTP3F, level = ('1'));
%freqdata3(order = 3, invar = dcct20);
%freqdata2(order = 4, invar1 = dcct20, invar2 = ANYSTP3F, level = ('1'));
%freqdata3(order = 5, invar = dcct30);
%freqdata2(order = 6, invar1 = dcct30, invar2 = ANYSTP3F, level = ('1'));
%freqdata3(order = 7, invar = dcct40);
%freqdata2(order = 9, invar1 = dcct40, invar2 = ANYSTP3F, level = ('1'));

%mend table2_datachunk;

%macro pdr_datachunk();

%freqdata3(order = 1, invar = dcct10);
%freqdata2(order = 2, invar1 = dcct10, invar2 = anypdr, level = ('1'));
%freqdata3(order = 3, invar = dcct20);
%freqdata2(order = 4, invar1 = dcct20, invar2 = anypdr, level = ('1'));
%freqdata3(order = 5, invar = dcct30);
%freqdata2(order = 6, invar1 = dcct30, invar2 = anypdr, level = ('1'));
%freqdata3(order = 7, invar = dcct40);
%freqdata2(order = 9, invar1 = dcct40, invar2 = anypdr, level = ('1'));

%mend pdr_datachunk;
```

```
%macro table3_datachunk();

%freqdata1(order = 1, invar = step3, level = ('1'));
%freqdata1(order = 2, invar = snpdr, level = ('1'));
%freqdata1(order = 3, invar = pdr, level = ('1'));

%freqdata1(order = 5, invar = photocoag, level = ('1'));


%mend table3_datachunk;



libname edic '/prj/niddk/ims_analysis/DCCT_EDIC/private_orig_data/edret18_responses/';

data edret18;
    set edic.edret18_v2;
               if fetdrspat =2 or fetdrspat = 3 then ma_only = 1;
               else ma_only = 0;
               if fetdrspat = 4 or fetdrspat = 5 then mildnpdr = 1;
               else mildnpdr = 0;
               if fetdrspat = 6 or fetdrspat = 7 or fetdrspat = 8 or fetdrspat = 9 then moderatenpdr = 1;
               else moderatenpdr=0;
               if anyfocal3=1 or anyscat3 =1 or ANYAVEGF=1 then photocoag =1;
               else photocoag =0;
               HBAMMOL = ((fhba1c - 2.152) * 10.931);
               HBAMMOL_dcct = ((dcct_hba - 2.152) * 10.931);
               HBAMMOL_edic = ((fedic_hba - 2.152) * 10.931);


data ret18;
  set edret18;
  where cycle=16.5;

proc sort data=ret18;
by mask_pat cycle dtedyear;

data ret18;
  set ret18;
  by mask_pat cycle dtedyear;
  if last.cycle then output ret18;

data csme18;
  set edret18;
  where csmepanal=1 and cycle=16.5;

proc sort data=csme18;
by mask_pat cycle dtedyear;

data csme18;
```

```
  set csme18;
  by mask_pat cycle dtedyear;
  if last.cycle;

data test;
       merge csme18
             ret18;
       by mask_pat;

data testpop;
       set test;
       keep mask_pat;

proc sort data=testpop;
by mask_pat;

data testonly;
  merge edret18 testpop(in=val1);
  by mask_pat;
  if val1 then output testonly;

proc sort data=testonly;
by mask_pat cycle dtedyear;

data testonly;
  set testonly;
  by mask_pat cycle dtedyear;
  if last.cycle then output testonly;



proc contents data=edret18;

proc format ;
       value $groupf 'EXPERIMENTAL' = 'INT'
                                                    'STANDARD' = 'CONV';

%let caser = group;
%let timeline = dtedyear = 0;

data accumfreq1 accummean1;
       set _null_;

data table1;
       set edret18;

proc sort data=table1;
       by &caser;

proc freq data=table1;
       title2 'Checking Case Counts for DCCT Baseline';
```

```
        tables &caser /nopercent nocum;
        where &timeline;
        format &caser $groupf.;

%datachunk();

data accumtab1 ;
        set accumfreq1 accummean1;
        if &caser = ' ' then delete;

proc sort data=accumtab1;
        by &caser orderer;

proc print data=accumtab1;
        title2 'Means and Frequencies for DCCT Baseline';
        var name charall orderer;
        by &caser;
        where orderer not in (28, 28.1, 30, 30.1, 46.1, 46.2, 46.3, 46.4, 46.5);
        format &caser $groupf.;

%let timeline = edicyear = 0;

data accumfreq1 accummean1;
        set _null_;

data table1;
        set edret18;

proc sort data=table1;
        by &caser;

proc freq data=table1;
        title2 'Checking Case Counts for End of DCCT';
        tables &caser /nopercent nocum;
        where &timeline;
        format &caser $groupf.;

%datachunk();

data accumtab2;
        set accumfreq1 accummean1;
                if &caser = ' ' then delete;


proc sort data=accumtab2;
        by &caser orderer;

proc print data=accumtab2;
        title2 'Means and Frequencies for End of DCCT';
        var name charall orderer;
        by &caser;
```

```
        where orderer not in  (26, 27, 30, 30.1, 42, 43, 44, 45, 46);
        format &caser $groupf.;

%let timeline = cycle=16.5;

data accumfreq1 accummean1;
        set _null_;

data table1;
        set testonly;


proc sort data=table1;
        by &caser;

proc freq data=table1;
        title2 'Checking Case Counts for EDIC years 15-18';
        tables &caser /nopercent nocum;
        where &timeline;
        format &caser $groupf.;

%datachunk();

data accumtab3;
        set accumfreq1 accummean1;

        if &caser = ' ' then delete;


proc sort data=accumtab3;
        by &caser orderer;

proc print data=accumtab3;
        title2 'Means and Frequencies for EDIC Years 15-18';
        var name charall orderer;
        by &caser;
        where orderer not in (26, 27, 28, 28.1, 42, 43, 44, 45, 46.5);
        format &caser $groupf.;

proc sort data=edret18;
        by mask_pat dtedyear;

data at_risk;
        set edret18;
        by mask_pat dtedyear;
        retain scat_v .;
        if first.mask_pat then scat_v = .;
        if anyascat = 1 and scat_v = . then scat_v = dtedyear;
        if last.mask_pat then output at_risk;

data outcomes;
```

```
        merge at_risk (in=val1 keep = mask_pat scat_v)
              edret18 (in=val2);
        by mask_pat;
        if val1 and val2 and prevanal =1 then output outcomes;

data further_outcome;
        set outcomes;
        by mask_pat dtedyear;
        if last.mask_pat and (ANSTP3FV>=100 or ANSTP3FV=.) and dtedyear>=100 and (scat_v=. or scat_v>=100) and step3f ne '.D' and
anystp3f ne '.D' then output further_outcome;

proc sort data=further_outcome;
        by &caser;

data pdr_outcome;
        set outcomes;
        by mask_pat;
        if last.mask_pat and (anypdrv>=100 or anypdrv=.) and dtedyear>=100 and (scat_v=. or scat_v>=100) then output pdr_outcome;

proc sort data=pdr_outcome;
        by &caser;

data accumfreq2 accumfreq3;
        set _null_;

data table1;
        set further_outcome;

proc sort data=table1;
        by &caser;

proc freq data=table1;
        title2 'Checking Case Counts for For All at Risk for Further 3 Step Progression at DCCT Closeout';
        tables &caser /nopercent;
        format &caser $groupf.;

proc freq data = table1 noprint;
        title2 'Checking Frequencies of Those with Further 3 Step Progression for all Retinopathy Levels';
        tables ANYSTP3F /out=counts;
        by &caser;
        format &caser $groupf.;

proc print data = counts;
        where ANYSTP3F = 1;
        format percent 8.1;

%table2_datachunk();

data accumtab4;
        set accumfreq2 accumfreq3;
        if &caser = ' ' then delete;
```

```
proc sort data=accumtab4;
       by &caser orderer;

proc print data=accumtab4;
       title2 'Counts and Frequencies of Those At Risk For and with Further than 3 Step Progression by Retinopathy Level';
       var name2 name charall orderer;
       by &caser;
       format &caser $groupf.;

data accumfreq2 accumfreq3;
       set _null_;

data table1;
       set pdr_outcome;

proc sort data=table1;
       by &caser;

proc freq data=table1;
       title2 'Checking Case Counts for For All at Risk for PDR at DCCT Closeout';
       tables &caser /nopercent;
       format &caser $groupf.;

proc freq data = table1;
       title2 'Checking Frequencies of Those with PDR for all Retinopathy Levels';
       tables anypdr /out=counts;
       by &caser;
       format &caser $groupf.;

proc print data = counts;
       where anypdr = 1;
       format percent 8.1;



%pdr_datachunk();

data accumtab7;
       set accumfreq2 accumfreq3;
       if &caser = ' ' then delete;


proc sort data=accumtab7;
       by &caser orderer;

proc print data=accumtab7;
       title2 'Counts and Frequencies of Those At Risk For and with PDR by Retinopathy Level';
       var name2 name charall orderer;
       by &caser;
       format &caser $groupf.;
```

```
proc sort data=edret18;
by mask_pat;

data csme18pop;
  set csme18;
  keep mask_pat;

proc sort data=csme18pop;
by mask_pat;

data csme18only;
  merge edret18 csme18pop(in=val1);
  by mask_pat;
  if val1 and csmepanal=1 then output csme18only;

proc sort data=csme18only;
by mask_pat cycle dtedyear;

data csme18only;
  set csme18only;
  by mask_pat cycle dtedyear;
  if last.cycle then output csme18only;

proc sort data=csme18only;
       by cycle;

proc freq data=csme18only;
       tables group*csme /list;
       where cycle in (0, 10, 16.5);
       by cycle;
       title3 'Table 3 -CSME numbers';

proc sort data=edret18;
by mask_pat;

data ret18pop;
  set ret18;
  keep mask_pat;

proc sort data=ret18pop;
by mask_pat;

data ret18only;
  merge edret18 ret18pop(in=val1);
  by mask_pat;
  if val1 then output ret18only;

proc sort data=ret18only;
by mask_pat cycle dtedyear;
```

```
data ret18only;
  set ret18only;
  by mask_pat cycle dtedyear;
  if last.cycle then output ret18only;

%let timeline = cycle  = 0;

data accumfreq1 ;
       set _null_;

data table1;
       set ret18only;


proc sort data=table1;
       by &caser;

proc freq data=table1;
       title2 'Checking Case Counts for DCCT Closeout';
       tables &caser /nopercent nocum missing;
       where &timeline;
       format &caser $groupf.;

%table3_datachunk();

data accumtab5;
       set accumfreq1 ;
       if &caser = ' ' then delete;

proc sort data=accumtab5;
       by &caser orderer;

proc print data=accumtab5;
       title2 'Retinopathy Complications After EDIC at DCCT Baseline';
       var name charall orderer;
       by &caser;
       format &caser $groupf.;


%let timeline = cycle  = 10;

data accumfreq1 ;
       set _null_;

data table1;
       set ret18only;


proc sort data=table1;
       by &caser;
```

```
proc freq data=table1;
       title2 'Checking Case Counts for EDIC year 10';
       tables &caser /nopercent nocum;
       where &timeline ;
       format &caser $groupf.;

%table3_datachunk();

data accumtab6;
       set accumfreq1 ;
       if &caser = ' ' then delete;


proc sort data=accumtab6;
       by &caser orderer;

proc print data=accumtab6;
       title2 'Retinopathy Complications After EDIC at EDIC year 10';
       var name charall orderer;
       by &caser;
       format &caser $groupf.;


%let timeline = cycle=16.5;

data accumfreq1 ;
       set _null_;

data table1;
       set ret18only;

proc sort data=table1;
       by &caser;

proc freq data=table1;
       title2 'Checking Case Counts for EDIC years 15-18';
       tables &caser /nopercent nocum;
       where &timeline ;
       format &caser $groupf.;

%table3_datachunk();

data accumtab6;
       set accumfreq1 ;
       if &caser = ' ' then delete;


proc sort data=accumtab6;
       by &caser orderer;

proc print data=accumtab6;
```

```
title2 'Retinopathy Complications After EDIC at EDIC years 15-18';
var name charall orderer;
by &caser;
format &caser $groupf.;
```