Dataset Integrity Check for the GFR (DCCT/EDIC) Data File

Prepared by Michael Spriggs
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton MD 20705
April 13, 2015

Table of 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 that was provided by the Data Coordinating Center (DCC) for this replication is located in the “EDIC\EDIC_Analysis_Datasets\SAS_DATA\” folder in the data package.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by The DCCT/EDIC Research Group in The New England Journal of Medicine, 2011 December 22. To verify the integrity of the datasets, two tables from the paper were checked (Tables B, D)

5 Results

Table 1 in the publication [1], Table 1. Demographic and Clinical Characteristics of the Participants at Baseline and at Closeout of the Diabetes Control and Complications Trial (DCCT) and at Year 16 of the Epidemiology of Diabetes Interventions and Complications (EDIC) Study, According to DCCT Treatment Group. 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 exact with only inconsequential discrepancies.

Table 2 in the publication [1], Table 2. Incidence of an Impaired Glomerular Filtration Rate (GFR) and Secondary Outcomes. Table C lists the variables that were used in the replication and Table D compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are exact.

6 Conclusions

The NIDDK repository is confident that the GFR data file to be distributed is a copy of the manuscript data.

7 References

[1] The DCCT/EDIC Research Group. Intensive Diabetes Therapy and Glomerular Filtration Rate in Type 1 Diabetes. The New England Journal of Medicine 2011;365(25):2366-2376. doi:10.1056/NEJMoa1111732.

Table A: Variables used to replicate Table 1. Demographic and Clinical Characteristics of the Participants at Baseline and at Closeout of the Diabetes Control and Complications Trial (DCCT) and at Year 16 of the Epidemiology of Diabetes Interventions and Complications (EDIC) Study, According to DCCT Treatment Group.

Characteristic	Variable(s)
Demographic characteristics: Age (yr)	Age0, Age99, ATTAGE
Demographic characteristics: Female sex (%)	Female
Medical history: Duration of diabetes (yr)	DURYR0, DURYR99 ATT_DUR
Medical history: DCCT primary cohort (%)‡	PRIMARY
Medical history: Hypertension (%)§	HT00, HT99, HT
Medical history: Hyperlipidemia (%)¶	HLIP00, HLIP99, HLIP
Medical history: Current smoking (%)	SMOKE00, SMOKE99, SMOKE
Medical history: Current alcohol use (%)	DRINK00, DRINK99, DRINK
Medical treatment: Insulin pump or ≥3 daily insulin injections	MDI
Medical treatment: Glucose monitoring ≥4 times/day	GLUC4
Medical treatment: Any	ANTIHYP
Medical treatment: ACE inhibitor or ARB	ACEARB
Physical examination findings: Body-mass index††	BMI00, BMI99, BMI
Physical examination findings: Systolic	SBP00, SBP99, SBP
Physical examination findings: Diastolic	DBP00, DBP99, DBP
Physical examination findings: Mean arterial pressure (mm Hg)	MBP00, MBP99, MBP
Laboratory values: Glycated hemoglobin (%)‡‡	HBAEL, DTMEANHB, EDIC_HBA
Laboratory values: Median (mg/24 hr)	AER
Laboratory values: Interquartile range (mg/24 hr)	AER
Laboratory values: ≥30 mg/24 hr (%)	AER30
Laboratory values: ≥300 mg/24 hr (%)	AER300
Laboratory values: Serum creatinine (mg/dl)	eSCR
Laboratory values: Total cholesterol	CHL00, CHL99, CHOL
Laboratory values: HDL cholesterol	HDL00, HDL99, HDL
Laboratory values: LDL cholesterol	LDL00, LDL99, LDL
Laboratory values: Triglycerides	TRG00, TRG99, TRIG

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

Characteristic	DCCT at Baseline Intensive Therapy (Manuscript N = 711	DCCT at Baseline Intensive Therapy (DSIC N=711)	DCCT at Baseline Intensive Therapy (DIFF N=0)
Demographic characteristics: Age (yr)	27.1±7.1	27.1±7.1	0,0
Demographic characteristics: Female sex (%)	48.5	48.5	0
Medical history: Duration of diabetes (yr)	6.0±4.2	6.0±4.2	0,0
Medical history: DCCT primary cohort (%)‡	49.0	49.0	0,0
Medical history: Hypertension (%)§	0	0	0
Medical history: Hyperlipidemia (%)¶	22.8	22.8	0
Medical history: Current smoking (%)	20.5	20.5	0
Medical history: Current alcohol use (%)	37.8	37.8	0
Medical treatment: Insulin pump or ≥3 daily insulin injections	0	0	0
Medical treatment: Glucose monitoring ≥4 times/day	0	0	0
Medical treatment: Any	0	0	0
Medical treatment: ACE inhibitor or ARB	0	0	0
Physical examination findings: Body-mass index††	23.4±2.7	23.4±2.7	0,0
Physical examination findings: Systolic	114.5±11.3	114.5±11.3	0,0
Physical examination findings: Diastolic	73.1±8.2	73.1±8.2	0,0
Physical examination findings: Mean arterial pressure (mm Hg)	86.9±8.2	86.9±8.2	0,0
Laboratory values: Glycated hemoglobin (%)‡‡	9.1±1.6	9.1±1.6	0,0
Laboratory values: Median (mg/24 hr)	11.5	11.5	0
Laboratory values: Interquartile range (mg/24 hr)	7.2–17.3	7.2–17.3	0-0
Laboratory values: ≥30 mg/24 hr (%)	11.7	11.7	0
Laboratory values: ≥300 mg/24 hr (%)	0	0	0
Laboratory values: Serum creatinine (mg/dl)	0.68±0.14	0.68±0.14	0,0
Laboratory values: Total cholesterol	177.1±32.8	177.1±32.8	0,0
Laboratory values: HDL cholesterol	50.8±12.3	50.8±12.3	0,0
Laboratory values: LDL cholesterol	110.3±28.7	110.3±28.7	0,0
Laboratory values: Triglycerides	80.8±43.3	80.8±43.3	0,0

Characteristic	DCCT at Baseline Conventional Therapy (Manuscript N = 730	DCCT at Baseline Conventional Therapy (DSIC N=730)	DCCT at Baseline Conventional Therapy (DIFF N=0)
Demographic characteristics: Age (yr)	26.5±7.1	26.5±7.1	0,0
Demographic characteristics: Female sex (%)	45.9	45.9	0
Medical history: Duration of diabetes (yr)	5.7±4.1	5.7±4.1	0,0
Medical history: DCCT primary cohort (%)‡	51.8	51.8	0
Medical history: Hypertension (%)§	0.3	0.3	0
Medical history: Hyperlipidemia (%)¶	23.3	23.3	0
Medical history: Current smoking (%)	21.6	21.6	0
Medical history: Current alcohol use (%)	39.9	39.9	0
Medical treatment: Insulin pump or ≥3 daily insulin injections	0	0	0
Medical treatment: Glucose monitoring ≥4 times/day	0	0	0
Medical treatment: Any	0	0	0
Medical treatment: ACE inhibitor or ARB	0	0	0
Physical examination findings: Body-mass index††	23.5±2.9	23.5±2.9	0,0
Physical examination findings: Systolic	114.6±11.4	114.6±11.4	0,0
Physical examination findings: Diastolic	72.9±8.7	72.9±8.7	0,0
Physical examination findings: Mean arterial pressure (mm Hg)	86.8±8.6	86.8±8.6	0,0
Laboratory values: Glycated hemoglobin (%)‡‡	9.1±1.6	9.1±1.6	0,0
Laboratory values: Median (mg/24 hr)	11.5	11.5	0
Laboratory values: Interquartile range (mg/24 hr)	7.2–18.7	7.2–18.7	0,0
Laboratory values: ≥30 mg/24 hr (%)	10.1	10.1	0
Laboratory values: ≥300 mg/24 hr (%)	0	0	0
Laboratory values: Serum creatinine (mg/dl)	0.68±0.14	0.68±0.14	0
Laboratory values: Total cholesterol	175.7±33.6	175.7±33.6	0,0
Laboratory values: HDL cholesterol	50.3±12.3	50.3±12.3	0,0
Laboratory values: LDL cholesterol	109.1±29.4	109.1±29.4	0,0
Laboratory values: Triglycerides	81.8±51.3	81.8±51.3	0,0

Characteristic	End of DCCT Intensive Therapy (Manuscript N = 698	End of DCCT Intensive Therapy (DSIC N=698)	End of DCCT Intensive Therapy (DIFF N=0)
Demographic characteristics: Age (yr)	33.4±7.0	33.4±7.0	0,0
Demographic characteristics: Female sex (%)	49.0	49.0	0
Medical history: Duration of diabetes (yr)	12.1±4.9	12.1±4.9	0,0
Medical history: DCCT primary cohort (%)‡	49.1	49.1	0
Medical history: Hypertension (%)§	0.7	0.7	0
Medical history: Hyperlipidemia (%)¶	26.0	26.0	0
Medical history: Current smoking (%)	23.1	23.1	0
Medical history: Current alcohol use (%)	36.3	36.3	0
Medical treatment: Insulin pump or ≥3 daily insulin injections	97.4	97.4	0
Medical treatment: Glucose monitoring ≥4 times/day	52.8	52.8	0
Medical treatment: Any	—	—	0
Medical treatment: ACE inhibitor or ARB	—	—	0
Physical examination findings: Body-mass index††	26.6±4.3	26.6±4.3	0,0
Physical examination findings: Systolic	116.6±11.5	116.6±11.5	0,0
Physical examination findings: Diastolic	74.8±8.7	74.8±8.7	0,0
Physical examination findings: Mean arterial pressure (mm Hg)	88.8±8.7	88.8±8.7	0,0
Laboratory values: Glycated hemoglobin (%)‡‡	7.3±0.9	7.3±0.9	0,0
Laboratory values: Median (mg/24 hr)	8.6	8.6	0
Laboratory values: Interquartile range (mg/24 hr)	5.8–14.1	5.8–14.4	0-,-0.3
Laboratory values: ≥30 mg/24 hr (%)	10.2	10.2	0
Laboratory values: ≥300 mg/24 hr (%)	1.4	1.4	0
Laboratory values: Serum creatinine (mg/dl)	0.73±0.14	0.73±0.14	0
Laboratory values: Total cholesterol	180.3±30.5	180.3±30.5	0,0
Laboratory values: HDL cholesterol	51.0±12.9	51.0±12.9	0,0
Laboratory values: LDL cholesterol	112.5±27.1	112.5±27.1	0,0
Laboratory values: Triglycerides	84.2±52.6	84.2±52.6	0,0

Characteristic	End of DCCT Conventional Therapy (Manuscript N = 717	End of DCCT Conventional Therapy (DSIC N=717)	End of DCCT Conventional Therapy (DIFF N=0)
Demographic characteristics: Age (yr)	32.8±7.0	32.8±7.0	0,0
Demographic characteristics: Female sex (%)	45.9	45.9	0
Medical history: Duration of diabetes (yr)	11.7±4.8	11.7±4.8	0,0
Medical history: DCCT primary cohort (%)‡	51.7	51.7	0
Medical history: Hypertension (%)§	1.8	1.8	0
Medical history: Hyperlipidemia (%)¶	29.7	29.7	0
Medical history: Current smoking (%)	23.2	23.2	0
Medical history: Current alcohol use (%)	38.9	38.9	0
Medical treatment: Insulin pump or ≥3 daily insulin injections	5.0‖	5.0‖	0
Medical treatment: Glucose monitoring ≥4 times/day	3.8‖	3.8‖	0
Medical treatment: Any	—	—	0
Medical treatment: ACE inhibitor or ARB	—	—	0
Physical examination findings: Body-mass index††	25.0±3.1‖	25.0±3.1‖	0,0
Physical examination findings: Systolic	116.6±11.9	116.6±11.9	0,0
Physical examination findings: Diastolic	74.4±8.9	74.4±8.9	0,0
Physical examination findings: Mean arterial pressure (mm Hg)	88.5±8.8	88.5±8.8	0,0
Laboratory values: Glycated hemoglobin (%)‡‡	9.1±1.3‖	9.1±1.3‖	0,0
Laboratory values: Median (mg/24 hr)	10.1‖	10.1‖	0
Laboratory values: Interquartile range (mg/24 hr)	5.8–20.2	5.8–20.2	0,0
Laboratory values: ≥30 mg/24 hr (%)	17.7‖	17.7‖	0
Laboratory values: ≥300 mg/24 hr (%)	3.2†	3.2†	0
Laboratory values: Serum creatinine (mg/dl)	0.72±0.18	0.72±0.18	0
Laboratory values: Total cholesterol	184.0±37.3	184.0±37.3	0,0
Laboratory values: HDL cholesterol	51.8±13.1	51.8±13.1	0,0
Laboratory values: LDL cholesterol	114.6±31.9	114.6±31.9	0,0
Laboratory values: Triglycerides	88.1±50.8†	88.1±50.8†	0,0

Characteristic	EDIC at Year 16 Intensive Therapy (Manuscript N = 618	EDIC at Year 16 Intensive Therapy (DSIC N=618)	EDIC at Year 16 Intensive Therapy (DIFF N=0)
Demographic characteristics: Age (yr)	50.4±6.9	50.4±6.9	0,0
Demographic characteristics: Female sex (%)	48.4	48.4	0
Medical history: Duration of diabetes (yr)	28.7±5.0	28.7±5.0	0,0
Medical history: DCCT primary cohort (%)‡	48.7	48.7	0
Medical history: Hypertension (%)§	53.7	53.7	0
Medical history: Hyperlipidemia (%)¶	65.7	65.7	0
Medical history: Current smoking (%)	13.3	13.3	0
Medical history: Current alcohol use (%)	42.9	42.9	0
Medical treatment: Insulin pump or ≥3 daily insulin injections	97.6	97.6	0
Medical treatment: Glucose monitoring ≥4 times/day	65.4	65.4	0
Medical treatment: Any	56.2	56.2	0
Medical treatment: ACE inhibitor or ARB	53.1	53.1	0
Physical examination findings: Body-mass index††	29.4±13.0	29.0±5.8	0.4+-7.2
Physical examination findings: Systolic	122.1±14.6	122.1±14.6	0,0
Physical examination findings: Diastolic	72.5±9.1	72.5±9.1	0,0
Physical examination findings: Mean arterial pressure (mm Hg)	89.0±9.6	89.0±9.6	0,0
Laboratory values: Glycated hemoglobin (%)‡‡	7.9±1.1	7.9±1.1	0,0
Laboratory values: Median (mg/24 hr)	11.5	11.5	0
Laboratory values: Interquartile range (mg/24 hr)	7.2–20.2	7.2–20.2	0,0
Laboratory values: ≥30 mg/24 hr (%)	19.4	19.0	-0.4
Laboratory values: ≥300 mg/24 hr (%)	3.2	3.2	0
Laboratory values: Serum creatinine (mg/dl)	0.85±0.33	0.85±0.33	0
Laboratory values: Total cholesterol	175.1±36.1	175.1±36.1	0,0
Laboratory values: HDL cholesterol	61.0±18.7	61.0±18.7	0,0
Laboratory values: LDL cholesterol	97.4±30.1	97.4±30.1	0,0
Laboratory values: Triglycerides	84.1±50.8	84.1±50.8	0,0

Characteristic	EDIC at Year 16 Conventional Therapy (Manuscript N = 604	EDIC at Year 16 Conventional Therapy (DSIC N=604)	EDIC at Year 16 Conventional Therapy (DIFF N=0)
Demographic characteristics: Age (yr)	49.4±6.9†	49.4±6.9†	0,0
Demographic characteristics: Female sex (%)	46.9	46.9	0
Medical history: Duration of diabetes (yr)	28.2±4.9	28.2±4.9	0,0
Medical history: DCCT primary cohort (%)‡	51.5	51.5	0
Medical history: Hypertension (%)§	51.2	51.2	0
Medical history: Hyperlipidemia (%)¶	65.2	65.2	0
Medical history: Current smoking (%)	11.8	11.8	0
Medical history: Current alcohol use (%)	44.7	44.7	0
Medical treatment: Insulin pump or ≥3 daily insulin injections	96.2	96.2	0
Medical treatment: Glucose monitoring ≥4 times/day	70.3	70.3	0
Medical treatment: Any	59.3	59.3	0
Medical treatment: ACE inhibitor or ARB	57.0	57.0	0
Physical examination findings: Body-mass index††	28.2±4.8	28.2±4.8	0
Physical examination findings: Systolic	121.2±15.2	121.2±15.2	0,0
Physical examination findings: Diastolic	72.2±8.8	72.2±8.8	0,0
Physical examination findings: Mean arterial pressure (mm Hg)	88.5±9.6	88.5±9.6	0,0
Laboratory values: Glycated hemoglobin (%)‡‡	8.0±1.0	8.0±1.0	0,0
Laboratory values: Median (mg/24 hr)	13.0†	13.0†	0
Laboratory values: Interquartile range (mg/24 hr)	7.2–28.8	7.2–28.8	0,0
Laboratory values: ≥30 mg/24 hr (%)	22.6	23.1	-0.5
Laboratory values: ≥300 mg/24 hr (%)	7.3‖	7.3‖	0
Laboratory values: Serum creatinine (mg/dl)	0.89±0.59	0.89±0.59	0
Laboratory values: Total cholesterol	172.2±37.4	172.2±37.4	0,0
Laboratory values: HDL cholesterol	60.6±17.5	60.6±17.5	0,0
Laboratory values: LDL cholesterol	94.9±30.1	94.9±30.2	0+--0.1
Laboratory values: Triglycerides	82.1±58.3†	82.1±58.3†	0,0

Table C: Variables used to replicate Table 2. Incidence of an Impaired Glomerular Filtration Rate (GFR) and Secondary Outcomes.

Characteristic	Variable(s)
Impaired GFR‡	ANYSCG60
Onset during DCCT	dtedyear
Onset during EDIC	dtedyear
Estimated GFR <45 ml/min/1.73 m_2	ANYCG45
Estimated GFR <30 ml/min/1.73 m_2§	ANYCG30
End-stage renal disease§	ANYESRD
Combined outcome of impaired GFR or death¶	E_COMP

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

Characteristic	Intensive Diabetes Therapy Manuscript	Intensive Diabetes Therapy DSIC	Intensive Diabetes Therapy DIFF
Impaired GFR‡	24	24	0
Onset during DCCT	1	1	0
Onset during EDIC	23	23	0
Estimated GFR <45 ml/min/1.73 m_2	24	24	0
Estimated GFR <30 ml/min/1.73 m_2§	13	13	0
End-stage renal disease§	8	8	0
Combined outcome of impaired GFR or death¶	53	53	0

Characteristic	Conventional Diabetes Therapy Manuscript	Conventional Diabetes Therapy DSIC	Conventional Diabetes Therapy DIFF
Impaired GFR‡	46	46	0
Onset during DCCT	3	3	0
Onset during EDIC	43	43	0
Estimated GFR <45 ml/min/1.73 m_2	39	39	0
Estimated GFR <30 ml/min/1.73 m_2§	23	23	0
End-stage renal disease§	16	16	0
Combined outcome of impaired GFR or death¶	80	80	0

```
libname gfr "/prj/niddk/ims_analysis/DCCT_EDIC/private_created_data/edic_new_data/GFR2010";

data gfr2010;
	set gfr.gfr2010;

%macro means_table1 (var1, var2, var3);

	proc means data = gfr2010 N MEAN STD maxdec=1;
	var &var1;
	where dtedyear = 0 and GROUP = "EXPERIMENTAL";
	title3 "Variable &var1, dtedyear = 0 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N MEAN STD maxdec=1;
	var &var1;
	where dtedyear = 0 and GROUP = "STANDARD";
	title3 "Variable &var1, dtedyear = 0 and GROUP = STANDARD";

	proc means data = gfr2010 N MEAN STD maxdec=1;
		var &var2;
		where dtedyear = 99 and GROUP = "EXPERIMENTAL";
		title3 "Variable &var2, dtedyear = 99 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N MEAN STD maxdec=1;
		var &var2;
		where dtedyear = 99 and GROUP = "STANDARD";
		title3 "Variable &var2, dtedyear = 99 and GROUP = STANDARD";

	proc means data = gfr2010 N MEAN STD maxdec=1;
		var &var3;
		where dtedyear = 1600 and GROUP = "EXPERIMENTAL";
		title3 "Variable &var3, dtedyear = 1600 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N MEAN STD maxdec=1;
		var &var3;
		where dtedyear = 1600 and GROUP = "STANDARD";
		title3 "Variable &var3, dtedyear = 1600 and GROUP = STANDARD";

%mend means_table1;

%macro means_table2 (var1, var2, var3);

	proc means data = gfr2010 N MEAN STD maxdec=2;
	var &var1;
	where dtedyear = 0 and GROUP = "EXPERIMENTAL";
	title3 "Variable &var1, dtedyear = 0 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N MEAN STD maxdec=2;
	var &var1;
	where dtedyear = 0 and GROUP = "STANDARD";
	title3 "Variable &var1, dtedyear = 0 and GROUP = STANDARD";

	proc means data = gfr2010 N MEAN STD maxdec=2;
		var &var2;
		where dtedyear = 99 and GROUP = "EXPERIMENTAL";
		title3 "Variable &var2, dtedyear = 99 and GROUP = EXPERIMENTAL";
```

```
	proc means data = gfr2010 N MEAN STD maxdec=2;
		var &var2;
		where dtedyear = 99 and GROUP = "STANDARD";
		title3 "Variable &var2, dtedyear = 99 and GROUP = STANDARD";

	proc means data = gfr2010 N MEAN STD maxdec=2;
		var &var3;
		where dtedyear = 1600 and GROUP = "EXPERIMENTAL";
		title3 "Variable &var3, dtedyear = 1600 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N MEAN STD maxdec=2;
		var &var3;
		where dtedyear = 1600 and GROUP = "STANDARD";
		title3 "Variable &var3, dtedyear = 1600 and GROUP = STANDARD";

%mend means_table2;

%macro median_table1 (var1, var2, var3);

	proc means data = gfr2010 N median P25 P75 maxdec=1;
	var &var1;
	where dtedyear = 0 and GROUP = "EXPERIMENTAL";
	title3 "Variable &var1, dtedyear = 0 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N median P25 P75 maxdec=1;
	var &var1;
	where dtedyear = 0 and GROUP = "STANDARD";
	title3 "Variable &var1, dtedyear = 0 and GROUP = STANDARD";

	proc means data = gfr2010 N median P25 P75 maxdec=1;
		var &var2;
		where dtedyear = 99 and GROUP = "EXPERIMENTAL";
		title3 "Variable &var2, dtedyear = 99 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N median P25 P75 maxdec=1;
		var &var2;
		where dtedyear = 99 and GROUP = "STANDARD";
		title3 "Variable &var2, dtedyear = 99 and GROUP = STANDARD";

	proc means data = gfr2010 N median P25 P75 maxdec=1;
		var &var3;
		where dtedyear = 1600 and GROUP = "EXPERIMENTAL";
		title3 "Variable &var3, dtedyear = 1600 and GROUP = EXPERIMENTAL";

	proc means data = gfr2010 N median P25 P75 maxdec=1;
		var &var3;
		where dtedyear = 1600 and GROUP = "STANDARD";
		title3 "Variable &var3, dtedyear = 1600 and GROUP = STANDARD";

%mend median_table1;


%macro freq_table1 (var1, var2, var3);

	proc freq data = gfr2010;
		table &var1*GROUP*dtedyear / list;
```

```
            where dtedyear = 0 and GROUP = "EXPERIMENTAL";
            title3 "Variable &var1, dtedyear = 0 and GROUP = EXPERIMENTAL";

        run;

        proc freq data = gfr2010;
            table &var1*GROUP*dtedyear / list;
            where dtedyear = 0 and GROUP = "STANDARD";
            title3 "Variable &var1, dtedyear = 0 and GROUP = STANDARD";

        run;

        proc freq data = gfr2010;
            table &var2*GROUP*dtedyear / list;
            where dtedyear = 99 and GROUP = "EXPERIMENTAL";
            title3 "Variable &var2, dtedyear = 99 and GROUP = EXPERIMENTAL";

        run;

        proc freq data = gfr2010;
            table &var2*GROUP*dtedyear / list;
            where dtedyear = 99 and GROUP = "STANDARD";
            title3 "Variable &var2, dtedyear = 99 and GROUP = STANDARD";

        run;

        proc freq data = gfr2010;
            table &var3*GROUP*dtedyear / list;
            where dtedyear = 1600 and GROUP = "EXPERIMENTAL";
            title3 "Variable &var3, dtedyear = 1600 and GROUP = EXPERIMENTAL";

        run;

        proc freq data = gfr2010;
            table &var3*GROUP*dtedyear / list;
            where dtedyear = 1600 and GROUP = "STANDARD";
            title3 "Variable &var3, dtedyear = 1600 and GROUP = STANDARD";

        run;


%mend freq_table1;

run;


%macro inc_rates (var1, var2);

        proc sort data = gfr2010_nodup;
            by group;


        proc summary data=gfr2010_nodup nway;
          var &var1 &var2;
          class group;
          output out=rates(drop=_type_ _freq_) sum=&var1 &var2;
```

```
        run;

        data rates;
          set rates;
          _rate=1000*(&var1/&var2);

        run;

        proc print data = rates;
          title 'Table of cases, person-years, and rates per 1000 person-years';
          var &var1 &var2 _rate ;
        run;

%mend inc_rates;


%means_table1(AGE0, AGE99, ATTAGE); run;

%freq_table1(female, female, female); run;

%means_table1(DURYR0, DURYR99, ATT_DUR); run;

%freq_table1(PRIMARY, PRIMARY, PRIMARY); run;

%freq_table1(HT00, HT99, HT); run;

%freq_table1(HLIP00, HLIP99, HLIP); run;

%freq_table1(SMOKE00, SMOKE99, SMOKE); run;

%freq_table1(DRINK00, DRINK99, DRINK); run;

%freq_table1(MDI, MDI, MDI); run;

%freq_table1(GLUC4, GLUC4, GLUC4); run;

%freq_table1(ANTIHYP, ANTIHYP, ANTIHYP); run;

%freq_table1(ACEARB, ACEARB, ACEARB); run;

%means_table1(BMI00, BMI99, BMI); run;

%means_table1(SBP00, SBP99, SBP); run;

%means_table1(DBP00, DBP99, DBP); run;

%means_table1(MBP00, MBP99, MBP); run;

%means_table1(HBAEL, DTMEANHB, EDIC_HBA); run;

%median_table1(AER, AER, AER); run;

%freq_table1(AER30, AER30, AER30); run;

%freq_table1(AER300, AER300, AER300); run;
```

```
%means_table2(eSCR, eSCR, eSCR); run;

%means_table1(CHL00, CHL99, CHOL); run;

%means_table1(HDL00, HDL99, HDL); run;

%means_table1(LDL00, LDL99, LDL); run;

%means_table1(TRG00, TRG99, TRIG); run;

proc sort data = gfr2010;
        by PATIENT;

data gfr2010_nodup;
        set gfr2010;
        by PATIENT;
        retain have_esrd have_impaired_gfr have_gfr_less_45 have_gfr_less_30 have_combine_outcome 0;
        if first.PATIENT then do;
                have_esrd = 0;
                have_impaired_gfr = 0;
                have_gfr_less_45 = 0;
                have_gfr_less_30 = 0;
                have_combine_outcome = 0;
        end;
        if ANYESRD = 1 then have_esrd = 1;
        if ANYSCG60 = 1 then have_impaired_gfr = 1;
        if ANYCG45 = 1 then have_gfr_less_45 = 1;
        if ANYCG30 = 1 then have_gfr_less_30 = 1;
        if E_COMP = 1 then have_combine_outcome = 1;

        if last.PATIENT then output gfr2010_nodup;

proc freq data = gfr2010_nodup;
        table
        have_impaired_gfr*group
        have_gfr_less_45*group
        have_gfr_less_30*group
        have_esrd*group
        have_combine_outcome*group
         / list missing;

data gfr_2010_detailed;
  merge gfr2010_nodup(keep=PATIENT have_impaired_gfr) gfr2010;
  by PATIENT;

data gfr_2010_short(drop=check_flag);
  set gfr_2010_detailed;
  by PATIENT;
  retain check_flag;
  length dtedyear_flag $4.;
  if 0<=dtedyear<=99 then dtedyear_flag="DCCT";
  else if 100<=dtedyear<=9999 then dtedyear_flag="EDIC";
  else do; abort; return; end;
  if first.patient then check_flag=0;
  if ANYSCG60=1 and check_flag=0 then do;
```

```
    check_flag=1;
    output;
  end;

proc freq data = gfr_2010_short;
        tables have_impaired_gfr*group*dtedyear_flag/missing list;
        where have_impaired_gfr = 1;
  title3 'Onset determination';
```