Dataset Integrity Check
EDIC 8-year Neuropathy Analysis Dataset

Sylvia E. Cohn, Principal Author
March 18, 2007

As a partial check of the integrity of the Neuropathy Analysis dataset archived in the NIDDK data repository, a number of analyses were performed to duplicate results published by Martin et al. in *Diabetes Care* in 2006.[1] Appendix A presents the full text of this article. Appendix B presents a comparison of the DCC's report on distributions of study variables to calculations made from the archived dataset, and Appendix C contains SAS code for the tabulations and analyses reported in this text and Appendix B.

The intent of this integrity check 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 on a first (or second) exercise 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 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. Thus, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses unless staff of the NIDDK Repository 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 (often in footnotes) those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

DCC SUMMARY STATISTICS.

The DCC for EDIC supplied summary statistics (Ns, percents, means, etc.) for the variables included in the analysis file. Appendix B includes the statistics provided by the DCC and three additional columns containing the values obtained in our dataset integrity check. Our analysis largely replicated these statistics[2] --- with the exception of the handling of MSNI missing values.[3] The DCC has informed us that this apparent discrepancy reflects their purposeful recoding of missing data codes to values of zero.

---

[1] Martin CL, Albers J, Herman WH, Cleary P, Waberski B, Greene DA, Stevens MJ, Feldman EL; DCCT/EDIC Research Group. Neuropathy among the diabetes control and complications trial cohort 8 years after trial completion. *Diabetes Care.* 2006 Feb;29(2):340-4.

[2] In addition to the handling of missing values described in Appendix D, the PT_Score variable in the archived dataset had 2,229 cases with a score of "3" (vs. 2,228 in DCC dataset documentation) and 298 cases with a score of "6" (vs. 297 in DCC dataset documentation).

[3] Across subjects and years of EDIC, we found 19 instances in which the *MSNI Checklist Score* (PT_SCORE) is coded as missing (.) but the binary indicator variable for "High" scores (PT_HIGH) is coded as zero ("not high", i.e., less than 7). Similarly, *MSNI Neuropathy Examination Scores* were calculated from 8 individual measurements made during the examination (EOCAL, EOCAR, EOC2L, EOC2R, EOC3L, EOC3R, EOC4L, EOC4R). These individual examination measurements are missing values for between 427 and 454 examinations -- and for 409 examinations, these measurements were missing for all 8 variables. The DCC codebook, however, indicates that there were only 349 missing values for the summary examination score.

COMPARISON TO PUBLISHED ARTICLE.

The published article reviewed evaluates the effect of prior intensive diabetes therapy on neuropathy. The following results, tables, graphs were obtained using a SAS analysis file extracted from the archived export file (*edicNEU8.xpt)* to reproduce the results in the published article. The published results are presented side by side with the values obtained from our dataset integrity check.

RESEARCH DESIGN AND METHODS

The publication reports that there were 1,398 EDIC subjects: 696 subjects from the intensive therapy group and 702 from the conventional therapy group, who had at least one MNSI assessment over the first 8 years of the study.

These numbers were verified. The analysis file contains one record per patient per EDIC year: 1,398 patients, for EDIC years 1-8, 10,543 observations in total.

RESULTS[4]

This paper presents the results of neuropathy status at the end of DCCT, at EDIC onset and up to 8 years into the EDIC study. The analysis investigates the association of neuropathy with original DCCT treatment group and with cumulative and concurrent glycemic levels. The paper also reports lower-extremity events (medical or surgical) associated with neuropathy.

Glycemic control in the EDIC study

Martin et al. begin their presentation by reporting on glycemic control over time in the treatment and control conditions (p. 341). As Table 1 shows, tabulations from the archived data closely match the values published by Martin et al. at DCCT completion, EDIC years 1, 5, and 8.

TABLE 1. Comparison of published HbA_{1c} values and those calculated from archived data by treatment group and time in study.

	PUBLISHED		DATASET CHECK	
TIME	Intensive	Conventional	Intensive	Conventional
At DCCT Completion	7.4	9.1	7.3	9.1
EDIC year 1	7.9	8.3	7.9	8.3
EDIC year 5	8.1	8.2	8.1	8.2
EDIC year 8	8.0	7.9	8.0	7.9

Neuropathy status at DCCT completion and EDIC onset

The published article notes that: "At completion of the DCCT, 19.1% of subjects fulfilled the DCCT criteria for definite clinical neuropathy (15.1% of the intensive therapy group and 23.0% of the conventional therapy group (p. 341)." Using the archived outcome variable DCCT_DN and the sample of 1,257 cited in published Table 1, we obtained identical results (19.1% overall, 15.1% for intensive treatment and 23.0% for conditional treatment. Similarly as shown in Table

[4] All our analyses use the MSNI summary scores and binary indicators in the archived dataset --- not variables recalculated from the individual MSNI items.

2, MSNI questionnaire and examination results obtained by tabulating the archived data are identical to those in the published article.

TABLE 2. Subjects satisfying MNSI criteria (questionnaire or examination) for neuropathy at the first annual EDIC study examination by treatment group [n (%)].

	Published		Dataset Check	
MSNI RESULTS	**Conventional**	**Intensive**	**Conventional**	**Intensive**
Ns	633	624	633	624
Positive Questionnaire	30 (4.7)	11 (1.8)	30 (4.7)	11 (1.8)
Positive Examination	177 (28.0)	111 (17.8)	177 (28.0)	111 (17.8)

The published article also reports that prior intense therapy reduced the odds of having symptoms or signs of neuropathy at the beginning of the EDIC. The results summarized in Table 3 indicate that analysis of the archived data yields results that are identical (within rounding error) to the published results.

TABLE 3. Impact of Intensive Therapy on reduction in odds of satisfying MSNI neuropathy criteria (questionnaire or examination) at first EDIC study measurement.

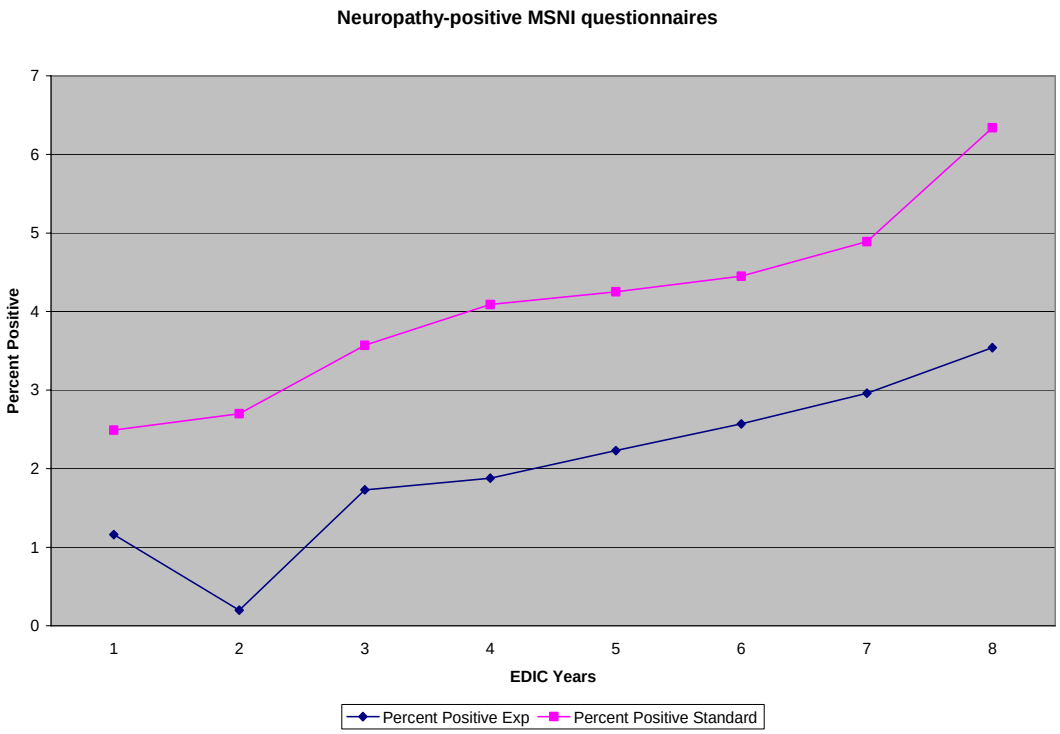
	Reduction in odds of Neuropathy	
MSNI RESULTS	**Published**	**Dataset Check**
Positive on Questionnaire	64% (27%-82%)	64% (27%-82%)
Positive on Examination	45% (27%-58%)	44% (27%-57%)

The published article also compares measures of neuropathy from DCCT and EDIC and reports that "nearly 20% of subjects without neuropathy at DCCT completion fulfilled MNSI examination criteria for neuropathy at first EDIC study evaluation". Our analysis of the archived dataset reproduced this result. Of the 1,238 subjects with no neuropathy at the end of the DCCT, we found that 19.14% fulfilled MSNI examination criteria for neuropathy.

Similarly, the published article reports that "at the first EDIC study evaluation, subjects classified with neuropathy at DCCT completion were 5 times more likely to have a positive MNSI questionnaire (9.7 vs. 1.8%, $p<0.0001$) and nearly twice as likely to have a positive MNSI examination (37.1 vs. 19.7%, $p<0.0001$) as subjects without neuropathy at DCCT completion". Our analysis yielded identical results: 9.7 vs. 1.8% ($p<0.0001$) for a positive MNSI questionnaire result and 37.1 vs. 19.7% ($p<0.0001$) for a positive MNSI examination result.

Persistence of DCCT treatment effect on neuropathy during the EDIC study

The frequency of neuropathy-positive MNSI questionnaires and MNSI examinations across 8 years of the EDIC study were presented in the paper in two figures. We plot below published results and results obtained from our analysis of the same time trends calculated from the archived data. It will be seen that there is high congruence between the published plots and those obtained from analysis of the archived data.



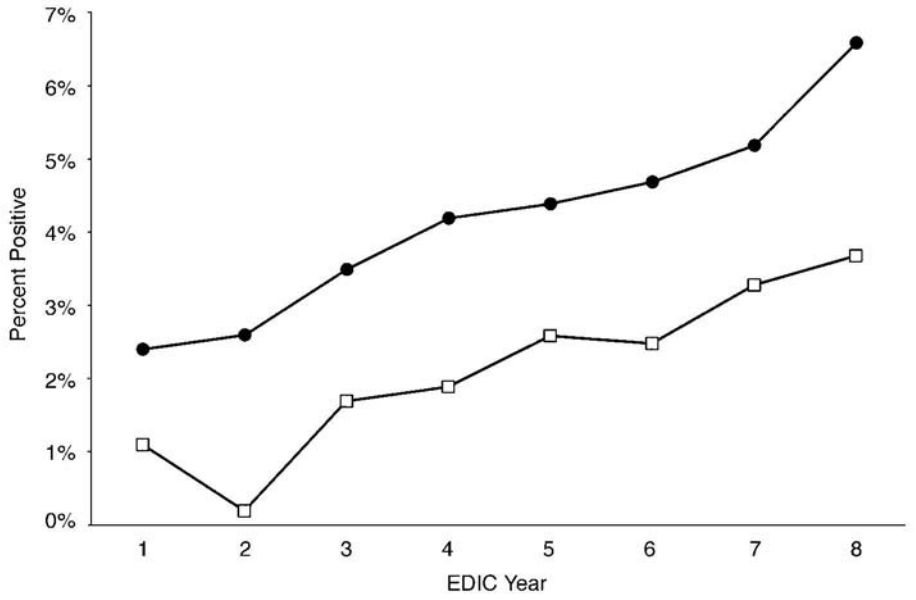
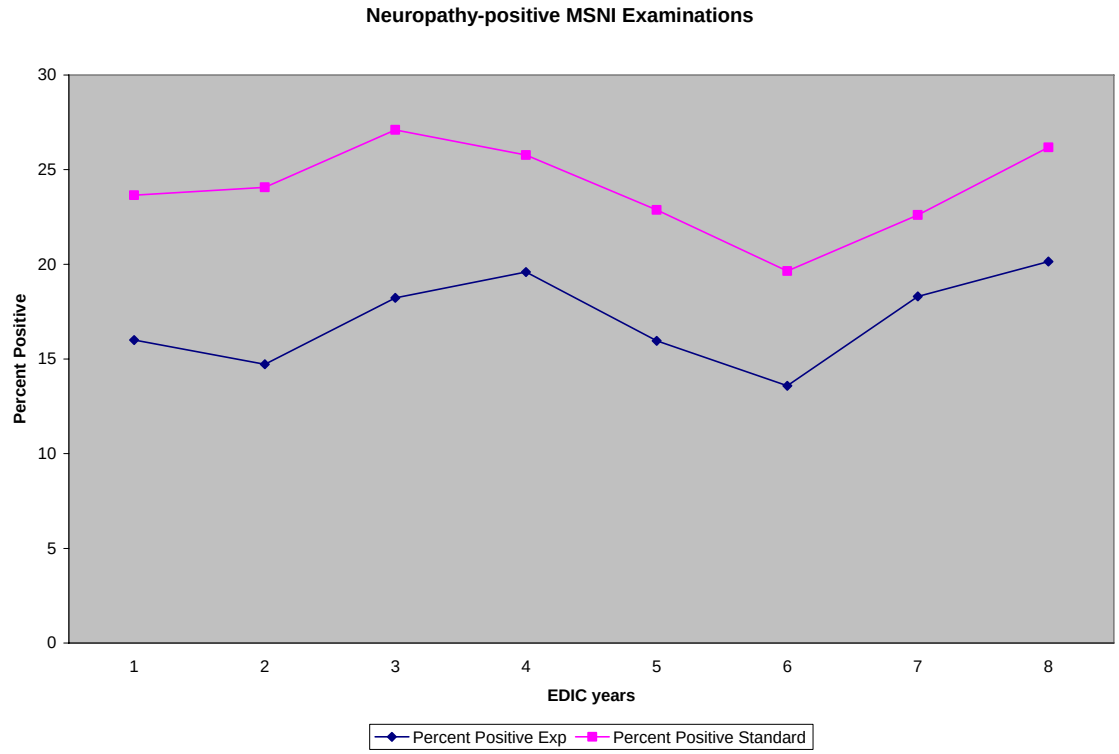


Figure 1—*Frequency of neuropathy-positive MNSI questionnaires across 8 years of the EDIC study among former DCCT conventional therapy (●) and intensive therapy (□) subjects without confirmed clinical neuropathy at the end of the DCCT. P < 0.0001 on average for all EDIC years combined.*

FIGURE 1. Trends over time by treatment group in the percent of subjects classified as positive on the MSNI neuropathy questionnaire: TOP: Calculated from Archived Data, BOTTOM: published.



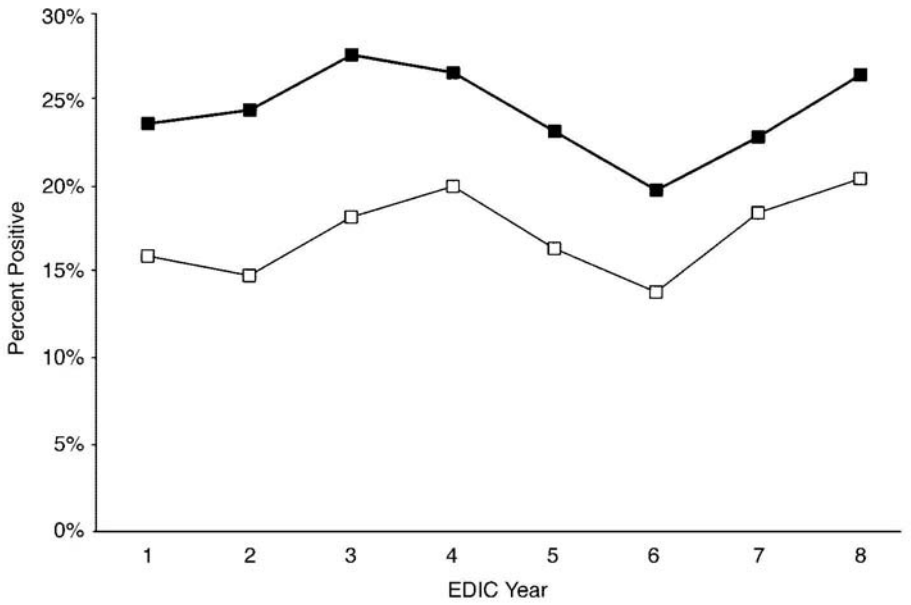


Figure 2—*Frequency of neuropathy-positive MNSI examinations across 8 years of the EDIC study among former DCCT conventional therapy (●) and intensive therapy (□) subjects without confirmed clinical neuropathy at the end of the DCCT. P < 0.0001 on average for all EDIC years combined.*

FIGURE 2. Trends over time by treatment group in the percent of subjects classified as positive on the MSNI neuropathy examination: TOP: Calculated from Archived Data, BOTTOM: published.

Lower extremities event associated with neuropathy

The paper also reports that during the study, 4 subjects from the intensive therapy group and 11 subjects from the conventional therapy group reported medical or surgical treatment. Also 2 subjects from the former intensive therapy group and five from the conventional therapy group underwent lower extremity amputations. We exactly reproduced the number of these events in our analyses of the archived data.

ATTACHMENT 1

Full Text of Article

Martin CL, Albers J, Herman WH, Cleary P, Waberski B, Greene DA, Stevens MJ, Feldman EL; DCCT/EDIC Research Group. Neuropathy among the diabetes control and complications trial cohort 8 years after trial completion. *Diabetes Care.* 2006 Feb;29(2):340-4.

APPENDIX B

EDIC 8-Year Neuropathy Analysis Dataset
Check of Summary Statistics

			DCC DOCUMENTATION			CALCULATED FROM ARCHIVED DATASET		
Variable	Variable Label	Category	n	%	Stat.	n	%	Stat.
Total	Total		10543	.	.	10543	.	.
C_HIGH/ RTI_C_HIGH	MNSI: CLINICAL SCORE > 2	. Missing	349	3.31	.	409	3.88	.
		0: No	7698	73.02	.	7638	72.45	.
		1: Yes	2496	23.67	.	2496	23.67	.
C_SCORE/ RTI_C_SCORE	MNSI: SCORE FROM CLINICAL EXAM	. Missing	349	3.31	.	409	3.88	.
		0	3761	35.67	.	3701	35.10	.
		0.5	275	2.61	.	275	2.61	.
		1	1215	11.52	.	1215	11.52	.
		1.5	225	2.13	.	225	2.13	.
		2	2222	21.08	.	2222	21.08	.
		2.5	315	2.99	.	315	2.99	.
		3	984	9.33	.	984	9.33	.
		3.5	140	1.33	.	140	1.33	.
		4	521	4.94	.	521	4.94	.
		4.5	98	0.93	.	98	0.93	.
		5	312	2.96	.	312	2.96	.
		5.5	41	0.39	.	41	0.39	.
		6	81	0.77	.	81	0.77	.
		7	2	0.02	.	2	0.02	.
		8	2	0.02	.	2	0.02	.
DCCT_DN	DCCT NEUROPATHY	.	167	1.58	.	167	1.58	.
		0: No	8433	79.99	.	8433	79.99	.
		1: Yes	1943	18.43	.	1943	18.43	.
DCCT_HBA	AVERAGE HBA1C DURING DCCT	Mean (N)	10543	.	8.1717	10543	.	8.1717
		STD	.	.	1.4090	.	.	1.4090
		Min	.	.	5.3950	.	.	5.3950
		Max	.	.	13.5750	.	.	13.5750
EDIC_HBA	AVERAGE HBA1C DURING EDIC (THRU CURRENT VISIT)	Mean (N)	10448	.	8.1681	10448	.	8.1681
		STD	.	.	1.2470	.	.	1.2470
		Min	.	.	4.4000	.	.	4.4000
		Max	.	.	15.1000	.	.	15.1000
EDICYEAR	EDIC YEAR	1	1341	12.72	.	1341	12.72	.
		2	1333	12.64	.	1333	12.64	.
		3	1329	12.61	.	1329	12.61	.
		4	1310	12.43	.	1310	12.43	.
		5	1308	12.41	.	1308	12.41	.
		6	1316	12.48	.	1316	12.48	.
		7	1304	12.37	.	1304	12.37	.
		8	1302	12.35	.	1302	12.35	.
EOB1	ARE YOUR LEGS AND/OR FEET NUMB	.	376	3.57	.	376	3.57	.

			DCC DOCUMENTATION			CALCULATED FROM ARCHIVED DATASET		
Variable	Variable Label	Category	n	%	Stat.	n	%	Stat.
		1: No	9351	88.69	.	9351	88.69	.
		2: Yes	816	7.74	.	816	7.74	.
EOB10	FEEL WEAK ALL OVER MOST OF THE TIME	.	372	3.53	.	372	3.53	.
		1: No	9911	94.01	.	9911	94.01	.
		2: Yes	260	2.47	.	260	2.47	.
EOB11	ARE YOUR SYMPTOMS WORSE AT NIGHT	.	415	3.94	.	415	3.94	.
		1: No	9347	88.66	.	9347	88.66	.
		2: Yes	781	7.41	.	781	7.41	.
EOB12	DO YOUR LEGS HURT WHEN YOU WALK	.	377	3.58	.	377	3.58	.
		1: No	9540	90.49	.	9540	90.49	.
		2: Yes	626	5.94	.	626	5.94	.
EOB13	ABLE TO SENSE YOUR FEET WHEN YOU WALK	.	380	3.60	.	380	3.60	.
		1: No	1180	11.19	.	1180	11.19	.
		2: Yes	8983	85.20	.	8983	85.20	.
EOB14	IS SKIN ON FEET SO DRY IT CRACKS OPEN	.	387	3.67	.	387	3.67	.
		1: No	8951	84.90	.	8951	84.90	.
		2: Yes	1205	11.43	.	1205	11.43	.
EOB15	HAVE YOU EVER HAD AN AMPUTATION	.	377	3.58	.	377	3.58	.
		1: No	10113	95.92	.	10113	95.92	.
		2: Yes	53	0.50	.	53	0.50	.
EOB2	EVER HAVE ANY BURNING PAIN IN LEGS/FEET	.	375	3.56	.	375	3.56	.
		1: No	9219	87.44	.	9219	87.44	.
		2: Yes	949	9.00	.	949	9.00	.
EOB3	ARE YOUR FEET TOO SENSITIVE TO TOUCH	.	377	3.58	.	377	3.58	.
		1: No	9817	93.11	.	9817	93.11	.
		2: Yes	349	3.31	.	349	3.31	.
EOB4	GET MUSCLE CRAMPS IN YOUR LEGS/FEET	.	386	3.66	.	386	3.66	.
		1: No	7618	72.26	.	7618	72.26	.
		2: Yes	2539	24.08	.	2539	24.08	.
EOB5	ANY PRICKLING FEELINGS IN LEGS/FEET	.	388	3.68	.	388	3.68	.
		1: No	8548	81.08	.	8548	81.08	.
		2: Yes	1607	15.24	.	1607	15.24	.
EOB6	HURT WHEN BED COVERS TOUCH YOUR SKIN	.	379	3.59	.	379	3.59	.
		1: No	9984	94.70	.	9984	94.70	.
		2: Yes	180	1.71	.	180	1.71	.
EOB7	ARE YOU ABLE TO TELL HOT/COLD WATER	.	373	3.54	.	373	3.54	.
		1: No	681	6.46	.	681	6.46	.
		2: Yes	9489	90.00	.	9489	90.00	.
EOB8	EVER HAD AN OPEN SORE ON YOUR FEET	.	382	3.62	.	382	3.62	.
EOB8	EVER HAD AN OPEN SORE ON YOUR FEET	1: No	8498	80.60	.	8498	80.60	.

			DCC DOCUMENTATION			CALCULATED FROM ARCHIVED DATASET		
Variable	Variable Label	Category	n	%	Stat.	n	%	Stat.
		2: Yes	1663	15.77	.	1663	15.77	.
EOB9	EVER HAD DIABETIC NEUROPATHY	.	402	3.81	.	402	3.81	.
		1: No	8554	81.13	.	8554	81.13	.
		2: Yes	1587	15.05	.	1587	15.05	.
EOCAL	APPEARANCE (L) NORMAL	.	433	4.11	.	433	4.11	.
		1: No	3102	29.42	.	3102	29.42	.
		2: Yes	7008	66.47	.	7008	66.47	.
EOCAR	APPEARANCE (R) NORMAL	.	427	4.05	.	427	4.05	.
		1: No	3104	29.44	.	3104	29.44	.
		2: Yes	7012	66.51	.	7012	66.51	.
EOCB1L	APPEARANCE (L) DEFORMITIES	.	10392	98.57	.	10392	98.57	.
		1: Yes	151	1.43	.	151	1.43	.
EOCB1R	APPEARANCE (R) DEFORMITIES	.	10394	98.59	.	10394	98.59	.
		1: Yes	149	1.41	.	149	1.41	.
EOCB2L	APPEARANCE (L) DRY SKIN, CALLUS	.	8226	78.02	.	8226	78.02	.
		1: Yes	2317	21.98	.	2317	21.98	.
EOCB2R	APPEARANCE (R) DRY SKIN, CALLUS	.	8194	77.72	.	8194	77.72	.
		1: Yes	2349	22.28	.	2349	22.28	.
EOCB3L	APPEARANCE (L) INFECTION	.	10272	97.43	.	10272	97.43	.
		1: Yes	271	2.57	.	271	2.57	.
EOCB3R	APPEARANCE (R) INFECTION	.	10268	97.39	.	10268	97.39	.
		1: Yes	275	2.61	.	275	2.61	.
EOCB4L	APPEARANCE (L) FISSURE	.	10394	98.59	.	10394	98.59	.
		1: Yes	149	1.41	.	149	1.41	.
EOCB4R	APPEARANCE (R) FISSURE	.	10386	98.51	.	10386	98.51	.
		1: Yes	157	1.49	.	157	1.49	.
EOCB5L	APPEARANCE (L) OTHER	.	9871	93.63	.	9871	93.63	.
		1: Yes	672	6.37	.	672	6.37	.
EOCB5R	APPEARANCE (R) OTHER	.	9867	93.59	.	9867	93.59	.
		1: Yes	676	6.41	.	676	6.41	.
EOC2L	ULCERATION (L)	.	453	4.30	.	453	4.30	.
		1: Absent	10029	95.12	.	10029	95.12	.
		2: Present	61	0.58	.	61	0.58	.
EOC2R	ULCERATION (R)	.	427	4.05	.	427	4.05	.
		1: Absent	10050	95.32	.	10050	95.32	.
		2: Present	66	0.63	.	66	0.63	.
EOC3L	ANKLE REFLEXES (L)	.	454	4.31	.	454	4.31	.
		1: Present	6325	59.99	.	6325	59.99	.
		2: Present/Reinforcement	1361	12.91	.	1361	12.91	.
		3: Absent	2403	22.79	.	2403	22.79	.
EOC3R	ANKLE REFLEXES (R)	.	452	4.29	.	452	4.29	.
		1: Present	6350	60.23	.	6350	60.23	.
		2: Present/Reinforcement	1395	13.23	.	1395	13.23	.

			DCC DOCUMENTATION			CALCULATED FROM ARCHIVED DATASET		
Variable	Variable Label	Category	n	%	Stat.	n	%	Stat.
		3: Absent	2346	22.25	.	2346	22.25	.
EOC4L	VIBRATION PERCEPTION AT GREAT TOE (L)	.	448	4.25	.	448	4.25	.
		1: Present	7649	72.55	.	7649	72.55	.
		2: Reduced	2185	20.72	.	2185	20.72	.
		3: Absent	261	2.48	.	261	2.48	.
EOC4R	VIBRATION PERCEPTION AT GREAT TOE (R)	.	445	4.22	.	445	4.22	.
		1: Present	7591	72.00	.	7591	72.00	.
		2: Reduced	2261	21.45	.	2261	21.45	.
		3: Absent	246	2.33	.	246	2.33	.
EOC5L	10 GRAM FILAMENT (L)	.	442	4.19	.	442	4.19	.
		1: Present	9457	89.70	.	9457	89.70	.
		2: Reduced	502	4.76	.	502	4.76	.
		3: Absent	142	1.35	.	142	1.35	.
EOC5R	10 GRAM FILAMENT (R)	.	443	4.20	.	443	4.20	.
		1: Present	9429	89.43	.	9429	89.43	.
		2: Reduced	537	5.09	.	537	5.09	.
		3: Absent	134	1.27	.	134	1.27	.
FSASDATE	DEIDENTIFIED FORMDATE	Mean (N)	10194	.	4115.762	10194	.	4115.762
		STD	.	.	1033.067	.	.	1033.067
		Min	.	.	1788.000	.	.	1788.000
		Max	.	.	6779.000	.	.	6779.000
HBA_DATE	DEIDENTIFIED HBA1C COLLECTION DATE	Mean (N)	10442	.	4098.925	10442	.	4098.925
		STD	.	.	1205.690	.	.	1205.690
		Min	.	.	-24818.0	.	.	-24818.0
		Max	.	.	13212.00	.	.	13212.00
HBAEL	HBA1C BASELINE ELIGIBILITY	Mean (N)	10543	.	9.0402	10543	.	9.0402
		STD	.	.	1.6039	.	.	1.6039
		Min	.	.	6.5600	.	.	6.5600
		Max	.	.	15.4200	.	.	15.4200
HBA1C	CURRENT HBA1C	Mean (N)	10448	.	8.1318	10448	.	8.1318
		STD	.	.	1.3768	.	.	1.3768
		Min	.	.	4.1000	.	.	4.1000
		Max	.	.	16.0000	.	.	16.0000
*PT_HIGH/ RTI_PT_HIGH	MNSI: PATIENT SCORE > 6	.	349	3.31	.	368	3.49	.
		0: No	9650	91.53	.	9631	91.35	.
		1: Yes	544	5.16	.	544	5.16	.
**PT_SCORE/ RTI_PT_SCORE	MNSI: SCORE FROM PATIENT CHECKLIST	.				368		
		0				147		
		1	560	5.31	.	560	5.31	.
		2	4473	42.43	.	4473	42.43	.
		3	2228	21.13	.	2229	21.13	.
		4	1225	11.62	.	1225	11.62	.
		5	600	5.69	.	600	5.69	.

			DCC DOCUMENTATION			CALCULATED FROM ARCHIVED DATASET		
Variable	Variable Label	Category	n	%	Stat.	n	%	Stat.
		6	398	3.78	.	397	3.78	.
		7	206	1.95	.	206	1.95	.
		8	124	1.18	.	124	1.18	.
		9	91	0.86	.	91	0.86	.
		10	55	0.52	.	55	0.52	.
		11	31	0.29	.	31	0.29	.
		12	19	0.18	.	19	0.18	.
		13	13	0.12	.	13	0.12	.
		14	4	0.04	.	4	0.04	.
		15	1	0.01	.	1	0.01	.
SUMAMP	AMPUTATION (EDIC TOTAL)	.: Missing	10488	99.48	.	10488	99.48	.
		1	55	0.52	.	55	0.52	.
SUMFOOT	FOOT ULCERS REQ MED/SURGICAL TREATMENT (EDIC TOTAL)	.: Missing	10440	99.02	.	10440	99.02	.
		1	86	0.82	.	86	0.82	.
		2	8	0.08	.	8	0.08	.
		3	9	0.09	.	9	0.09	.
SUMLEG	LEG ULCERS REQ MED/SURGICAL TREATMENT (EDIC TOTAL)	.: Missing	10528	99.86	.	10528	99.86	.
		1	15	0.14	.	15	0.14	.

APPENDIX C

SAS Programs Used in Dataset Check

TABLES

```
OPTIONS NOFMTERR;
libname data "c:\projects\NIDDK\EDIC\NewData\data";

data neuro_8yr;
   set data.neuro_8yr;


****APPENDIX 3: SUMMARY STATISTICS OF ALL VARIABLES;

   ****  EOB1--EOB15 1 IS NEGATIVE- 2 IS POSITIVE
   ****  NEOB1--NEOB15   0 IS NEGATIVE AND 1 IS POSITIVE;

DO I=1 TO 15;


ARRAY MSNI1 (15)  EOB1 EOB2 EOB3 EOB4 EOB5 EOB6 EOB7 EOB8 EOB9 EOB10 EOB11
                  EOB12 EOB13 EOB14 EOB15;

       MSNI1(I) = MSNI1(I)-1;
    END;


****RECREATING PT_SCORE;
RTI_PT_SCORE = SUM(EOB1,EOB2,EOB3,EOB4,EOB5,EOB6,EOB7,EOB8,
                  EOB9,EOB10,EOB11,EOB12,EOB13,EOB14,EOB15);

****REPRODUCING PT_HIGH FROM PT_SCORE;
IF RTI_PT_SCORE > 6 THEN RTI_PT_HIGH=1;
   ELSE IF 0<=RTI_PT_SCORE<=6 THEN RTI_PT_HIGH=0;
   ELSE RTI_PT_HIGH=.;

****RECREATING C_SCORE;
IF SUM(EOCAL,EOCAR,EOC2L,EOC2R,EOC3L,EOC3R,EOC4L,EOC4R)=. THEN RTI_C_SCORE=.;
   ELSE RTI_C_SCORE= ROUND(SUM(
(EOCAL = 1),
(EOCAR = 1),
(EOC2L = 2),
(EOC2R = 2),
(0.5 * (EOC3L = 2) + (EOC3L = 3)),
(0.5 * (EOC3R = 2) + (EOC3R = 3)),
(0.5 * (EOC4L = 2) + (EOC4L = 3)),
(0.5 * (EOC4R = 2) + (EOC4R = 3))));

****REPRODUCING C_HIGH FROM C_SCORE;
IF RTI_C_SCORE >2 then RTI_C_HIGH=1;
   ELSE IF RTI_C_SCORE = . THEN RTI_C_HIGH=.;
   ELSE IF RTI_C_SCORE <=2 THEN RTI_C_HIGH=0;


LABEL
  DCCT_DN  ='DCCT NEUROPATHY'
  DCCT_HBA ='AVERAGE HBA1C DURING DCCT'
  EDIC_HBA ='AVERAGE HBA1C DURING EDIC'
  HBAEL='HBA1C BASELINE ELIGIBILITY'
  HBA1C='CURRENT HBA1C';

****N=10543;

TITLE 'APPENDIX 3: SUMMARY STATISTICS OF ALL VARIABLES';


PROC FREQ; TABLES  C_SCORE RTI_C_SCORE C_HIGH RTI_C_HIGH/MISSING;

PROC FREQ; TABLES DCCT_DN EDICYEAR/MISSING;
```

```
PROC MEANS; VAR DCCT_HBA EDIC_HBA;

PROC FREQ;
   TABLES NEOB1 NEOB2 NEOB3 NEOB4 NEOB5 NEOB6 NEOB7 NEOB8
          NEOB9 NEOB10 NEOB11 NEOB12 NEOB13 NEOB14 NEOB15/MISSING;

PROC FREQ;
   TABLES EOCAL EOCAR EOCBIL EOCBIR EOCB2L EOCB2R EOCB3L EOCB3R EOCB4L EOCB4R
          EOCB5L EOCB5R EOC2L EOC2R EOC3L EOC3R EOC4L EOC4R EOC5L EOC5R/MISSING;

PROC MEANS;
   VAR FSASDATE HBA_DATE HBAEL HBA1C;

PROC FREQ; TABLES PT_HIGH RTI_PT_HIGH PT_SCORE RTI_PT_SCORE /MISSING;

PROC FREQ; TABLES SUMAMP SUMFOOT SUMLEG/MISSING;

RUN;
*****************************************************************************;

*******************REPRODUCE TABLES AND NUMBERS FROM MANUSCRIPT*****************;

**1- RESEARCH DESIGN AND METHODS************************************************
* THERE WERE 1,398 EDIC SUBJECTS: 696 EXPERIMENTAL (INTENSIVE TX) AND 702
* STANDARD TX WITH AT LEAST ONE ASSESSMENT IN 8 YRS
*****************************************************************************;


PROC SORT DATA=NEURO_8YR;
   BY MASK_PAT;

****AT LEAST ONE ASSESSMENT IN 8 YEARS;
DATA NEURO;
   SET NEURO_8YR;
   BY MASK_PAT;
   IF LAST.MASK_PAT;
RUN;

PROC FREQ; TABLES GROUP;
RUN;

*****************************************************************************
**2- OUTCOME MEASURES (page 341);
*****************************************************************************

PROC FREQ DATA=NEURO_8YR;
   TABLES PT_HIGH*RTI_PT_HIGH/MISSING;

PROC FREQ DATA=NEURO_8YR;
   TABLES C_HIGH*RTI_C_HIGH/MISSING;



**3- GLYCEMIC CONTROL IN THE EDIC STUDY (page 341)******************;
**TRYING TO REPRODUCE VALUES AT DCCT COMPLETION (7.4% AND 9.1% P<0.01)************
**RTI gets 7.3 and 9.1
*****************************************************************************;


DATA YEAR1;
   SET NEURO_8YR;
   IF EDICYEAR=1;
RUN;

OPTIONS LS=132 PAGENO=1;;
PROC SORT DATA=YEAR1; BY GROUP;
PROC TTEST DATA=YEAR1; VAR DCCT_HBA ; CLASS GROUP;
RUN;
```

```
**REPRODUCING VALUES AT 1, 5 AND 8 YEARS OF EDIC*******;

PROC SORT DATA=NEURO_8YR;
   BY GROUP;
PROC TTEST; VAR HBA1C; CLASS GROUP; WHERE EDICYEAR=1 ;
TITLE 'AIC AT EDIC - YEAR 1';

PROC SORT DATA=NEURO_8YR;
   BY GROUP;
PROC TTEST; VAR HBA1C; CLASS GROUP; WHERE EDICYEAR=5 ;
TITLE 'AIC AT EDIC - YEAR 5';

PROC SORT DATA=NEURO_8YR;
   BY GROUP;
PROC TTEST; VAR HBA1C; CLASS GROUP; WHERE EDICYEAR=8 ;
TITLE 'AIC AT EDIC - YEAR 8';
RUN;




***4- RESULTS************************************;

***1- NEUROPATHY STATUS AT DCCT COMPLETION AND EDIC ONSET;
***N=1257 (633 and 624);

DATA YEAR1A;
   SET NEURO_8YR;
   IF EDICYEAR=1;
   IF GROUP='EXPERIMENTAL' THEN NGROUP=1;
     ELSE IF GROUP='STANDARD' THEN NGROUP=0;
IF PT_SCORE NE . OR C_SCORE NE .; *n=1257;

PROC FREQ DATA=YEAR1A;
   TABLES DCCT_DN*NGROUP;
   TITLE 'NEUROPATHY STATUS AT COMPLETION OF DCCT';
RUN;


*****************************************************************************
* BEGINNING OF EDIC STUDY 3.3% + ACCORDING TO MSNI QUESTIONNAIRE FOR NEUROPATHY
* (1.8 AND 4.7)                                                             *
* AND 22.9 + ACCORDING TO MSNI EXAMINATION (17.8% AND 28.0%)                *
* THIS IS TABLE 1                                                           *
*****************************************************************************;

PROC FREQ DATA=YEAR1A;
   TABLES PT_HIGH*NGROUP;
   TITLE 'EDIC STUDY YEAR 1: NEUROPATHY ACCORDING TO MNSI QUESTIONNAIRE';
RUN;

PROC FREQ DATA=YEAR1A;
   TABLES C_HIGH*NGROUP;
   TITLE 'EDIC STUDY YEAR 1: NEUROPATHY ACCORDING TO MNSI EXAMINATION';

RUN;



*****************************************************************************
* PRIOR INTENSIVE THERAPY REDUCED ODDS OF HAVING SYMTOMS (MSNI QUESTIONNAIRE)    *
* AT BEGINNING OF EDIC BY 64% (95% CI 27-82%, P=0.0044) AND SIGNS OF NEUROPATHY  *
* BY 45% (27%-58%, P<0.0001.                                                *
* *************************************************************************;

PROC FREQ; TABLES NGROUP*PT_HIGH/CMH;
TITLE 'ODDS OF HAVING SYMPTOMS';
```

```
run;

PROC FREQ; TABLES NGROUP*C_HIGH/CMH;
TITLE 'ODDS OF HAVING SIGNS';
run;


*****************************************************************************
* MNSI questionnaire criteria for neuropathy                                *
* Case-Control      Mantel-Haenszel        0.3607       0.1791       0.7263  *
*                                                                           *
* OR 0.36  (0.18-0.73) for intensive therapy compared with standard         *
* ---> 64% decrease (27%-82%) - ok!!!  p=0.003 (not 0.0044)                 *
*                                                                           *
* MNSI symtoms for neuropathy                                               *
* Case-Control      Mantel-Haenszel        0.5574       0.4261       0.7293  *
* ---> 45% decrease (27%-58%) - ok!!!  p < 0.0001                           *
*****************************************************************************




*****************************************************************************
* 20% of subjects w/o neuropathy at DCCT show neuropathy at first EDIC exam *
* OK!!                                                                      *
*****************************************************************************;

DATA YEAR1A;
   SET NEURO_8YR;
   IF EDICYEAR=1;
   IF GROUP='EXPERIMENTAL' THEN NGROUP=1;
     ELSE IF GROUP='STANDARD' THEN NGROUP=0;
IF PT_SCORE NE . OR C_SCORE NE .; *n=1257;
PROC FREQ; TABLES DCCT_DN*C_HIGH;
TITLE 'TABLE WITH EDIC VARIABLES';
run;

*****************************************************************************
*MANUSCRIPT:
* SUBJECTS DCCT_DN=1  5 TIMES MORE LIKELY TO HAVE + QUESTIONNAIRE (9.7 VS 1.8%  *
*                     P <0.0001                                             *
*                                                                           *
*                     2 TIMES MORE LIKELY TO HAVE + EXAMINATION (37.1 VS 19.7%  *
*                     P <0.0001                                             *
*****************************************************************************;

PROC FREQ DATA=YEAR1A; TABLES DCCT_DN*PT_HIGH/CMH;
TITLE 'TABLE WITH EDIC VARIABLES';
run;

PROC FREQ DATA=YEAR1A; TABLES DCCT_DN*C_HIGH/CMH;
TITLE 'TABLE WITH EDIC VARIABLES';
run;

*****************************************************************************
*results   :                                                                *
* subjects dcct_dn=1  questionnaire or=5.87 (9.7 vs 1.8%)  p <.0001         *
*                                                                           *
*                                                                           *
*                     examination or=2.4  (37.1 vs 19.7%)  p <.0001         *
*                                                                           *
*****************************************************************************;


*****************************************************************************
* PERSISTENCE OF THE DCCT TREATMENT EFFECT ON NEUROPATHY OVER 8 YEARS EDIC STUDY *
*****************************************************************************;
```

```
****VALUES FOR GRAPHIC;

DATA NEUROPATHY;
   SET NEURO_8YR;
   IF DCCT_DN=0;
PROC SORT; BY EDICYEAR;
PROC FREQ; TABLES GROUP*PT_HIGH; BY EDICYEAR;
RUN;

PROC FREQ; TABLES GROUP*C_HIGH; BY EDICYEAR;
RUN;

*******************************************************************************
* values for graphic- do not see dip in neuropathy-positive msni questionnaire    *
* for experimental treatment at 6 yrs-                                            *
*                                                                                 *
* questionaire:                                                                   *
*                                                                                 *
* experimental    1.16      0.2     1.73    1.88    2.23    2.57    2.96    3.54  *
* standard        2.49      2.70    3.57    4.09    4.25    4.45    4.89    6.34  *
*                                                                                 *
* examination:                                                                    *
*                                                                                 *
* experimental    16.0      14.72   18.23   19.59   15.96   13.58   18.30   20.15 *
* standard        23.65     24.07   27.10   25.77   22.87   19.64   22.61   26.18 *
*******************************************************************************;



****************************************
**MANUSCRIPT- LOWER EXTREMITY EVENTS ASSOCIATED WITH
NEUROPATHY**********************************
* 15 SUBJECTS REPORTED TOTAL 22 LOWER EXTREMITIES ULCERS (20 FOOT AND 2 LEG ULCERS

* FEWER SUBJECTS IN THE DCCT INTENSIVE THERAPY GROUP DEVELOPED FOOT OR LEG ULCERS THAN
SUBJECTS
* IN CONVENTIONAL THERAPY GROUP 4 VS 11 P=0.01.
*
* 7 SUBJECTS UNDERWEN LOWER EXTREMITIES AMPUTATIONS 2 FROM THE FORMER THERAPY
INTENSIVE GROUP
* AND 5 IN THE CONVENTIONAL THERAPY GROUP 2 VS 5 P=0.45



data ulcers;
   set neuro_8yr;
   if sumfoot ne . or sumleg ne .;
proc sort;
   by mask_pat;

data ulcers1;
   set ulcers;
   by mask_pat;
   if first.mask_pat;
proc freq; tables sumfoot sumleg;
run;
proc print; var mask_pat sumfoot sumleg;
run;

*******************************************************************************
****NOTE:
*****# of lower extremity events- numbers match the 15 22 and (22 2) from the manuscript
*******************************************************************************;

data abn;
   set neuro_8yr;
proc sort data=abn;
   by mask_pat;
```

```
data abn1;
   set abn;
   by mask_pat;
   if first.mask_pat;
   if sumfoot ne . or sumleg ne . then ulcers=1;
      else ulcers=0;
   if sumamp = . then sumamp=0;
proc freq;
   tables group*ulcers/chisq;
   title 'lower extremity eventsby DCCT tx';
run;

***#s are 4 vs 11 but p=0.072 and not 0.01;

proc freq data=abn1;
   tables group*sumamp/chisq;
   title 'amputations';
run;


****************************************
******RESULTS:
****#s are 2 vs 5 but p=0.26 and not 0.;
****************************************:

PROC PRINT;  WHERE C_SCORE=0 AND RTI_C_SCORE=.;
  VAR EOCAL EOCAR EOC2L EOC2R EOC3L EOC3R EOC4L EOC4R;
  title 'check the raw values for discordant results';

GENMOD ANALYSIS

***OVERALL DIFFERENCES OVER 8 YEARS OF EDIC;


*****all these were attempts to try to reproduce the sections in the manuscript
"Persistence of the DCCT treatment effect on neuropathy during the EDIC study"
and
"Influence of cumulative and concurrent glycemic control on neuropathy"
;

*****Only the figures were included in the report-
*******************************************************************************
* MANUSCRIPT: MSNI QUESTIONNAIRE REDUCED 51% 95CI 30%-66%  P<0.0001             *
*             MSNI EXAMINATION REDUCED 43% 95CI 33%-52%    P<0.0001             *
*******************************************************************************;

OPTIONS LS=132 PAGENO=1;

PROC GENMOD DATA=NEUROPATHY DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = GROUP  /DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE 'OVERALL DIFFERENCES IN MSNI QUESTIONNAIRE OVER 8 YEARS OF EDIC';
RUN;

PROC GENMOD DATA=NEUROPATHY DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL C_HIGH = GROUP/DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH  COVB CORRW;
   TITLE 'OVERALL DIFFERENCES IN MSNI EXAMINATION OVER 8 YEARS OF EDIC';
RUN;


data OR_QUEST;

beta=-0.7309;
beta1=-1.207;
beta2=-0.2546;
```

```
OR=EXP(beta);
OR95LL=EXP(BETA1);
OR95HL=EXP(beta2);

proc print; var OR OR95LL OR95HL;
   TITLE 'OR AND 95% CI FOR DIFFERENCES IN MSNI QUESTIONNAIRE OVER 8 YEARS EDIC';
run;



data OR_EXAM;

beta=-0.445;
beta1=-0.635;
beta2=-0.25;

OR=EXP(beta);
OR95LL=EXP(BETA1);
OR95HL=EXP(beta2);

proc print; var OR OR95LL OR95HL;
   TITLE 'OR AND 95% CI FOR DIFFERENCES IN MSNI EXAMINATION OVER 8 YEARS EDIC';
run;
```

the likelihood of neuropathy based on the MSNI questionnaire and the MSNI examination was reduced 51% (95% CI 30-66%, p<0.0001) and 43% (33-52%, p<0.0001) respectively, among subjects with prior intensive therapy compared with conventional therapy across 8 years of EDIC follow-up

```
*******************************************************************************
* results:                                                                    *
*                                                                             *
* questionnaire-                                                              *
*                                    standard   95% confidence                *
* parameter              estimate    error       limits         z pr > |z|    *
*                                                                             *
* intercept               -3.1191   0.1528  -3.4185  -2.8196  -20.41   <.0001    *
* group     experimental  -0.7301   0.2430  -1.2064  -0.2537   -3.00   0.0027    *
* group     standard       0.0000   0.0000   0.0000   0.0000     .      .         *
*
* or = 0.48186  95% ci 0.29927    0.77592
*                                                                             *
* examination-                                                                *
*                                                                             *
*                                   standard   95% confidence                 *
* parameter              estimate    error       limits            z pr > |z|    *
*                                                                             *
* intercept               -1.1538   0.0657  -1.2825  -1.0251  -17.57   <.0001    *
* group     experimental  -0.4406   0.0968  -0.6304  -0.2509   -4.55   <.0001    *
* group     standard       0.0000   0.0000   0.0000   0.0000                      *
*
* or = 0.64365  95% ci 0.53238    0.77810
*******************************************************************************;

*******************************************************************************
* MANUSCRIPT: MSNI QUESTIONNAIRE REDUCED 51% 95CI 30%-66%  P<0.0001           *
*             MSNI EXAMINATION REDUCED 43% 95CI 33%-52%    P<0.0001           *
*******************************************************************************;

***************?????????DIFFERENCES?????????????????????????????????????????;
```

MANUSCRIPT- INFLUENCE OF CUMULATIVE AND CONCURRENT GLYCEMIC CONTROL ON NEUROPATHY*********

```
* NEUROPATHY (QUESTIONNAIRE AND EXAMINATION) ASSOCIATED WITH CUMULATIVE MEAN A1C LEVEL
*
* A 1% LOWER CUMULATIVE MEAN A1C REDUCED THE ODDS OF FULFILLING MNSI QUESTIONNAIRE
CRITERIA       *
* FOR NEUROPATHY BY 38% (95% CI 38-47%, P<0.0001) AND MNSI EXAMINATION CRITERIA
*
* FOR NEUROPATHY BY 27% (95% CI 22-32%, P<0.0001)
*
* NO SIGNIFICANT ASSOCIATION FOUND BETWEEN CONCURRENT A1C AND EITHER POSITIVE MNSI
QUESTIONNAIRE  *
* OR EXAMINATION
*
********************************************************************************
************;

OPTIONS NOCENTER PAGENO=1 LS=132;
TITLE1 'INFLUENCE OF CUMULATIVE AND CONCURRENT GLYCEMIC CONTROL ON NEUROPATHY';

PROC GENMOD DATA=NEURO_8YR DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = GROUP HBA1C EDIC_HBA/DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE2 'MODEL INCLUDES: TREATMENT, CUMULATIVE AND CONCURRENT GLYCEMIC CONTROL';
RUN;
PROC GENMOD DATA=NEURO_8YR DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = GROUP HBA1C /DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE2 'MODEL INCLUDES: TREATMENT AND CONCURRENT GLYCEMIC CONTROL';
RUN;
PROC GENMOD DATA=NEURO_8YR DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = GROUP EDIC_HBA/DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE2 'MODEL INCLUDES: TREATMENT AND CUMULATIVE GLYCEMIC CONTROL';
RUN;
PROC GENMOD DATA=NEURO_8YR DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = HBA1C EDIC_HBA/DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE2 'MODEL INCLUDES: CUMULATIVE AND CONCURRENT GLYCEMIC CONTROL';
RUN;
PROC GENMOD DATA=NEURO_8YR DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = EDIC_HBA/DIST=BIN LINK=LOGIT;
   REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE2 'MODEL INCLUDES:  CUMULATIVE GLYCEMIC CONTROL';
RUN;
PROC GENMOD DATA=NEURO_8YR DESCENDING;
   CLASS MASK_PAT GROUP;
   MODEL PT_HIGH = HBA1C DIST=BIN LINK=LOGIT;
REPEATED SUBJECT=MASK_PAT/TYPE=EXCH /*COVB CORRW*/;
   TITLE2 'MODEL INCLUDES: CONCURRENT GLYCEMIC CONTROL';
RUN;


***NOTE: manuscript says no significant association found between concurrent a1c and
either positive mnsi questionnaire*
*        or examination. i found association- doing year by year- alone in model or
after adjusting by prior treatment *
*        also overall - with treatment and without there is significant association
*
*        manuscript results: a 1% lower cumulative mean a1c reduced the odds
*
                             for positive questionnaire by 38% (95% ci 28-47%,
p<0.0001)                                *
                             for positive examination by 27% (95% ci 22-32%, p<0.0001)
*
```

```
* equivalent to get odds ratio of 0.62 (95% 0.53-0.72) and 0.73 (0.68-0.78)  can't
reproduce                               *
*************************************************************************
*
```