

Dataset Integrity Check for the
Nonalcoholic Steatohepatitis
(NASH) Pioglitazone vs Vitamin E vs
Placebo for Treatment of Non-Diabetic
Patients With Nonalcoholic
Steatohepatitis (PIVENS) Analysis Files

Prepared by Jane Wang
IMS Inc.

3901 Calverton Blvd, Suite 200 Calverton MD 20705

January 16, 2014

Table of Contents

1 Standard Disclaimer.....	1
2 Study Background.....	1
3 Archived Datasets.....	2
4 Statistical Methods.....	2
5 Results.....	3
6 Conclusions.....	4
7 References.....	4
Attachment A: SAS Code.....	12
Table A: Variables used to replicate Table 1: Baseline Characteristics of study subjects	5
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	6
Table C: Variables used to replicate Table 2: Primary Outcome and Changes in Histologic Features of the Liver after 96 Weeks of Treatment	8
Table D: Comparison of values computed in integrity check to reference article Table 2 values.....	9
Table E: Variables used to replicate Table 2: Changes in Serum Biochemical Level, Metabolic Factors, and Quality of Life from Baseline to 96 Weeks.....	10
Table F: Comparison of values computed in integrity check to reference article Table 3 values.....	11

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

Nonalcoholic steatohepatitis (NASH) is a common liver disease associated with obesity and diabetes. NASH is a progressive disorder that can lead to cirrhosis and liver failure, and there is no definitive treatment. Insulin resistance and oxidative stress (resulting in lipid peroxidation) are considered to be the two most important mechanisms in the pathogenesis of NASH.

PIVENS hypothesized that pioglitazone and vitamin E will lead to improvement in hepatic histology in nondiabetic adults with biopsy proven NASH through changes in insulin resistance or oxidative stress. Before and post-treatment liver biopsies were read centrally in a masked fashion for an assessment of steatohepatitis and a NAFLD Activity Score (NAS) consisting of steatosis, lobular inflammation, and hepatocyte ballooning.

PIVENS enrollment started in January 2005 and ended in January 2007 with 247 patients randomized to receive either pioglitazone (30 mg q.d.), vitamin E (800 IU q.d.), or placebo for 96 weeks. Participants were followed for an additional 24 weeks after stopping the treatment. Both at the initial evaluation and at the completion of treatment 96 weeks later, subjects underwent an assessment of body weight, height, and waist and hip circumferences, and blood samples were obtained for routine biochemical tests and assessment of fasting levels of lipids, glucose, and insulin. Body composition was assessed with the use of dual-energy x-ray absorptiometry. The Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) was administered for the assessment of quality of life.

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the NASH “NASHCRN_Data_Sharing_PIVENS_NEJM2010” folder in the data package. For this replication, variables were taken from “table1”, “table 2”, and “table 3” datasets. These datasets were analysis datasets created by the DCC.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Arun et al [1] in the New England Journal of Medicine in May 2010.

To verify the integrity of the three dataset, descriptive statistics of baseline characteristics were computed, by treatment group (Table B, Table D and Table E).

5 Results

Table 1 in the publication [1], Baseline Characteristics of the Study Subjects. Table A lists the variables that were used in the replication and Table B compare the results calculated from the archived data file to the results published in Table 1. The results of the replication are similar to published results, within rounding error.

Table 2 in the publication [1] Primary Outcome and Changes in Histologic Features of the Liver after 96 Weeks of Treatment. 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 2. The results of the replication are the same as published results.

Table 3 in the publication [1] Changes in Serum Biochemical Levels, Metabolic Factors, and Quality of Life from Baseline to 96 Weeks. Table E lists the variables that were used in the replication and Table F compares the results calculated from the archived data file to the results published in Table 3. The results of the replication are similar to published results, within rounding error.

6 Conclusions

The NIDDK repository is confident that the NASH Pivens data files to be distributed are a true copy of the study data.

7 References

1. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, Neuschwander-Tetri BA, Lavine JE, Tonascia J, Ünalp A, Van Natta M, Clark J, Brunt EM, Kleiner DE, Hoofnagle JH, and Robuck PR; for the NASH CRN. Pioglitazone, Vitamin E, or Placebo for Nonalcoholic Steatohepatitis. *New England Journal of Medicine*. 2010; 362:1675-85. PMID: PMC2928471 (MS# 2009-1)

Table A: Variables used to replicate Table 1: Baseline Characteristics of Study Subjects

Table Variable	Variables Used in Replication from the "Table 1" Dataset
Age (yr)	AGE
Female sex (%)	FEMALE
Hispanic	HISP
Nonwhite	WHITE
SF-36, physical component	PCS
SF-36, mental component	MCS
Alanine aminotransferase (U/liter)	ALT
Aspartate aminotransferase (U/liter)	AST
Glutamyltransferase (U/liter)	GGT
Alkaline phosphatase (U/liter)	ALKA
Total bilirubin (mg/dl)	BILIT
Triglycerides (mg/dl)	TRIG
Cholesterol (mg/dl)Total	CHOL
Cholesterol (mg/dl) High-density lipoprotein	HDL
Cholesterol (mg/dl)Low-density lipoprotein	LDL
Fasting serum glucose (mg/dl)	GLUC
Insulin resistance	HOMA
Weight (kg)	WEIGHT
Body-mass index	BMI
Waist circumference (cm)	WAISTCM
Body composition (% fat)	TOTF
Total NAFLD activity score	BLNAS
Steatosis	BLSTEATO
Lobular inflammation	BLINFLAM
Hepatocellular ballooning	BLBALL
Fibrosis stage	BLFIBRO
Absence of ballooning on central review (%)	BLBALLO

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

characteristic	Placebo (N=83) [Manuscript]	Placebo (N=83) [DSIC]	Placebo (N=83) [Difference]	Vitamin E (N=84) [Manuscript]	Vitamin E (N=84) [DSIC]	Vitamin E (N=84) [Difference]
Age (yr)	45.4 ,11.2	45.4 , 11.2	0.0 , 0.0	46.6 ,12.1	46.6 , 12.1	0.0 , 0.0
Female sex (%)	58	58	0	62	62	0
Hispanic	7	7	0	19	19	0
Nonwhite	11	11	0	15	15	0
SF-36, physical component	47 ,11	47 , 11	0 , 0	49 ,10	49 , 10	0 , 0
SF-36, mental component	47 ,12	47 , 10	0 , 2	49 ,10	49 , 11	0 , -1
Alanine aminotransferase (U/liter)	81 ,48	81 , 48	0 , 0	86 ,52	86 , 52	0 , 0
Aspartate aminotransferase (U/liter)	55 ,30	55 , 30	0 , 0	59 ,33	59 , 33	0 , 0
??-Glutamyltransferase (U/liter)	69 ,79	69 , 79	0 , 0	56 ,42	56 , 42	0 , 0
Alkaline phosphatase (U/liter)	82 ,25	82 , 25	0 , 0	77 ,25	77 , 25	0 , 0
Total bilirubin (mg/dl)	0.76 ,0.35	0.76 , 0.35	0.00 , 0.00	0.75 ,0.38	0.75 , 0.38	0.00 , 0.00
Triglycerides (mg/dl)	165 ,89	165 , 89	0 , 0	166 ,104	166 , 104	0 , 0
Cholesterol (mg/dl)Total	199 ,40	199 , 40	0 , 0	195 ,40	195 , 40	0 , 0
Cholesterol (mg/dl) High-density lipoprotein	43 ,11	43 , 11	0 , 0	44 ,13	44 , 13	0 , 0
Cholesterol (mg/dl)Low-density lipoprotein	125 ,35	125 , 35	0 , 0	119 ,35	119 , 35	0 , 0
Fasting serum glucose (mg/dl)	95 ,14	95 , 14	0 , 0	95 ,14	95 , 14	0 , 0
Insulin resistance??	5.5 ,5.1	5.5 , 5.1	0.0 , 0.0	5.2 ,4.0	5.2 , 4.0	0.0 , 0.0
Weight (kg)	99 ,21	99 , 21	0 , 0	94 ,24	95 , 24	-1 , 0
Body-mass index??	35 ,7	35 , 7	0 , 0	34 ,7	34 , 7	0 , 0
Waist circumference (cm)	109 ,14	109 , 14	0 , 0	107 ,15	107 , 15	0 , 0
Body composition (% fat)	40 ,9	40 , 9	0 , 0	39 ,9	39 , 9	0 , 0
Total NAFLD activity score	4.8 ,1.4	4.8 , 1.4	0.0 , 0.0	5.1 ,1.4	5.1 , 1.5	0.0 , -0.1
Steatosis	1.9 ,0.8	1.9 , 0.8	0.0 , 0.0	1.9 ,0.9	1.9 , 0.9	0.0 , 0.0
Lobular inflammation	1.6 ,.0.7	1.6 , 0.7	0.0 , 0.0	1.8 ,0.7	1.8 , 0.7	0.0 , 0.0
Hepatocellular ballooning	1.3 ,0.7	1.3 , 0.7	0.0 , 0.0	1.3 ,0.8	1.3 , 0.8	0.0 , 0.0
Fibrosis stage**	1.6 ,1.1	1.6 , 1.1	0.0 , 0.0	1.5 ,1.0	1.5 , 1.0	0.0 , 0.0
Absence of ballooning on central review (%)	17	17	0	18	18	0

characteristic	Pioglitazone (N=80) [Manuscript]	Pioglitazone (N=80) [DSIC]	Pioglitazone (N=80) [Difference]	Total (N=247) [Manuscript]	Total (N=247) [DSIC]	Total (N=247) [Difference]
Age (yr)	47.0 ,12.6	47.0 , 12.6	0.0 , 0.0	46.3 ,11.9	46.3 , 11.9	0.0 , 0.0
Female sex (%)	59	59	0	60	60	0
Hispanic	19	19	0	15	15	0
Nonwhite	8	8	0	12	11	1
SF-36, physical component	49 ,9	49 , 9	0 , 0	48 ,10	48 , 10	0 , 0
SF-36, mental component	49 ,8	49 , 10	0 , -2	48 ,10	48 , 10	0 , 0
Alanine aminotransferase (U/liter)	82 ,45	82 , 45	0 , 0	83 ,49	83 , 49	0 , 0
Aspartate aminotransferase (U/liter)	54 ,26	54 , 26	0 , 0	56 ,30	56 , 30	0 , 0
??-Glutamyltransferase (U/liter)	60 ,63	60 , 63	0 , 0	61 ,63	61 , 63	0 , 0
Alkaline phosphatase (U/liter)	86 ,39	86 , 39	0 , 0	81 ,30	81 , 31	0 , -1
Total bilirubin (mg/dl)	0.77 ,0.38	0.77 , 0.38	0.00 , 0.00	0.76 ,0.37	0.76 , 0.37	0.00 , 0.00
Triglycerides (mg/dl)	162 ,84	162 , 84	0 , 0	165 ,93	165 , 93	0 , 0
Cholesterol (mg/dl)Total	195 ,37	195 , 37	0 , 0	196 ,39	196 , 39	0 , 0
Cholesterol (mg/dl) High-density lipoprotein	45 ,12	45 , 12	0 , 0	44 ,12	44 , 12	0 , 0
Cholesterol (mg/dl)Low-density lipoprotein	120 ,31	120 , 31	0 , 0	122 ,34	122 , 34	0 , 0
Fasting serum glucose (mg/dl)	92 ,12	92 , 12	0 , 0	94 ,13	94 , 13	0 , 0
Insulin resistance??	5.0 ,3.8	5.0 , 3.8	0.0 , 0.0	5.2 ,4.3	5.2 , 4.3	0.0 , 0.0
Weight (kg)	97 ,23	97 , 23	0 , 0	97 ,23	97 , 23	0 , 0
Body-mass index??	34 ,6	34 , 6	0 , 0	34 ,7	34 , 7	0 , 0
Waist circumference (cm)	108 ,14	108 , 14	0 , 0	108 ,14	108 , 14	0 , 0
Body composition (% fat)	40 ,8	40 , 8	0 , 0	39 ,9	39 , 9	0 , 0
Total NAFLD activity score	5.0 ,1.4	5.0 , 1.4	0.0 , 0.0	4.9 ,1.4	4.9 , 1.4	0.0 , 0.0
Steatosis	2.0 ,0.8	2.0 , 0.8	0.0 , 0.0	1.9 ,0.8	1.9 , 0.8	0.0 , 0.0
Lobular inflammation	1.8 ,0.7	1.8 , 0.7	0.0 , 0.0	1.7 ,0.7	1.8 , 0.7	-0.1 , 0.0
Hepatocellular ballooning	1.1 ,0.8	1.1 , 0.8	0.0 , 0.0	1.3 ,0.8	1.3 , 0.8	0.0 , 0.0
Fibrosis stage**	1.4 ,0.9	1.4 , 0.9	0.0 , 0.0	1.5 ,1.0	1.5 , 1.0	0.0 , 0.0
Absence of ballooning on central review (%)	28	28	0	21	21	0

Table C: Variables used to replicate Table 2: Primary Outcome and Changes in Histologic Features of the Liver after 96 Weeks of Treatment

Table Variable	Variables Used in Replication from the "Table 2" Dataset
No. of subjects randomly assigned	IMPROVE
Subjects with improvement (%)	IMPROVE
No. of subjects with biopsy specimens at baseline and 96 wk	RESOLVE
Steatosis Subjects with improvement (%)	ISTEATO
Steatosis Mean change in score	CSTEATO
Lobular inflammation Subjects with improvement (%)	IINFLAM
Lobular inflammation Mean change in score	CINFLAM
Hepatocellular ballooning Subjects with improvement (%)	IBALL
Hepatocellular ballooning Mean change in score	CBALL
Total NAFLD activity score (mean change)	CNAS
Fibrosis Subjects with improvement (%)	IFIBRO
Fibrosis Mean change in score	CFIBRO
Resolution of definite nonalcoholic steatohepatitis(% of subjects)	RESOLVE

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

Characteristic	Placebo [Manuscript]	Placebo [DSIC]	Placebo [Difference]	Vitamin E [Manuscript]	Vitamin E [DSIC]	Vitamin E [Difference]	Pioglitazone [Manuscript]	Pioglitazone [DSIC]	Pioglitazone [Difference]
No. of subjects randomly assigned	83	83	0	84	84	0	80	80	0
Subjects with improvement (%)	19	19	0	43	43	0	34	34	0
No. of subjects with biopsy specimens at base	72	72	0	80	80	0	70	70	0
Steatosis Subjects with improvement (%)	31	31	0	54	54	0	69	69	0
Steatosis Mean change in score	-0.1	-0.1	0	-0.7	-0.7	0	-0.8	-0.8	0
Lobular inflammation Subjects with improvement	35	35	0	54	54	0	60	60	0
Lobular inflammation Mean change in score	-0.2	-0.2	0	-0.6	-0.6	0	-0.7	-0.7	0
Hepatocellular ballooning Subjects with impro	29	29	0	50	50	0	44	44	0
Hepatocellular ballooning Mean change in scor	-0.2	-0.2	0	-0.5	-0.5	0	-0.4	-0.4	0
Total NAFLD activity score (mean change)	-0.5	-0.5	0	-1.9	-1.9	0	-1.9	-1.9	0
Fibrosis Subjects with improvement (%)	31	31	0	41	41	0	44	44	0
Fibrosis Mean change in score	-0.1	-0.1	0	-0.3	-0.3	0	-0.4	-0.4	0
Resolution of definite nonalcoholic steatohep	21	21	0	36	36	0	47	47	0

Table E: Variables used to replicate Table 2: Changes in Serum Biochemical Level, Metabolic Factors, and Quality of Life from Baseline to 96 Weeks.

Table Variable	Variables Used in Replication from the "Table 3" Dataset
Alanine aminotransferase (U/liter)	CALT
Aspartate aminotransferase (U/liter)	CAST
?-Glutamyltransferase (U/liter)	CGGT
Alkaline phosphatase (U/liter)	CALKA
Total bilirubin (mg/dl)	CBILIT
Lipids Triglycerides (mg/dl)	CTRIG
Cholesterol (mg/dl) Total	CCHOL
Cholesterol (mg/dl) High-density lipoprotein	CHDL
Cholesterol (mg/dl) Low-density lipoprotein	CLDL
Fasting serum glucose (mg/dl)	CGLUC
Insulin resistance?	CHOMA
Weight (kg)	CWEIGHT
Body-mass index	CBMI
Waist circumference (cm)	CWAIST
Body composition (% fat)	CTOTF
SF-36 score, physical component	CPCS
SF-36 score, mental component	CMCS

Table F: Comparison of values computed in integrity check to reference article Table 3 values.

Characteristic	Placebo [Manuscript]	Placebo [DSIC]	Placebo [Difference]	Vitamin E [Manuscript]	Vitamin E [DSIC]	Vitamin E [Difference]	Pioglitazone [Manuscript]	Pioglitazone [DSIC]	Pioglitazone [Difference]
Alanine aminotransferase (U/liter)	-20.1	-20.1	0	-37	-37	0	-40.8	-40.8	0
Aspartate aminotransferase (U/liter)	-3.8	-3.8	0	-21.3	-21.3	0	-20.4	-20.4	0
Glutamyltransferase (U/liter)	-4	-4	0	-14	-14	0	-21.1	-21.1	0
Alkaline phosphatase (U/liter)	-3.8	-3.8	0	-9.3	-9.3	0	-12	-12	0
Total bilirubin (mg/dl)	0.06	0.06	0	0.04	0.04	0	-0.04	-0.04	0
Lipids Triglycerides (mg/dl)	-6.7	-6.7	0	-0.6	-0.6	0	-19.8	-19.8	0
Cholesterol (mg/dl) Total	-9.6	-9.6	0	-13.6	-13.6	0	-11.4	-11.5	0.1
Cholesterol (mg/dl) High-density lipoprotein	-1.9	-1.9	0	-0.9	-0.9	0	1.1	1.1	0
Cholesterol (mg/dl) Low-density lipoprotein	-5.8	-5.8	0	-12	-12	0	-8.1	-8.1	0
Fasting serum glucose (mg/dl)	1.8	1.8	0	1.8	1.8	0	-3.1	-3.1	0
Insulin resistance?	0.4	-0.2	0.6	0.4	1.4	-1	-0.7	-0.5	-0.2
Weight (kg)	0.7	0.7	0	0.4	0.3	0.1	4.7	4.7	0
Body-mass index	0.4	0.4	0	0.1	0.1	0	1.8	1.8	0
Waist circumference (cm)	0.2	0.2	0	-0.4	-0.4	0	3.3	3.3	0
Body composition (% fat)	0	0	0	0.4	0.4	0	2.7	2.7	0
SF-36 score, physical component	-0.3	-0.3	0	0.4	0.5	-0.1	-0.9	-0.9	0
SF-36 score, mental component	0.4	0.4	0	-0.5	-0.5	0	-1.9	-1.9	0

Attachment A: SAS Code

```
/******  
*****  
***Program: /prj/niddk/ims_analysis/NASH/prog_initial_analysis/nash_integrity_check.sas;  
***Programmer: Jane Wang  
***Date Created: 11/05/2013  
***Purpose: To perform a Dataset Integrity Check (DSIC) between the CIR data and the primary  
***outcome paper:  
***Pioglitazone, Vitamin E, or Placebo for Nonalcoholic Steatohepatitis Arun J. Sanyal, M.D.,  
***Naga Chalasani, M.B., B.S., Kris V. Kowdley, M.D., Arthur McCullough, M.D., Anna Mae Diehl,  
***M.D., Nathan M. Bass, M.D., Ph.D., Brent A. Neuschwander-Tetri, M.D., Joel E. Lavine, M.D.,  
*** Ph.D., James Tonascia, Ph.D., Aynur Unalp, M.D., Ph.D., Mark Van Natta, M.H.S., Jeanne Clark,  
***M.D., M.P.H., Elizabeth M. Brunt, M.D., David E. Kleiner, M.D., Ph.D., Jay H. Hoofnagle, M.D.,  
***and Patricia R. Robuck, Ph.D., M.P.H., for the NASH CRN  
***The numbers in Tables 1,2,and 3 of the primary outcome paper will compared to the NASH data  
***received;  
*****  
*****/;  
  
title1 "%sysfunc(getoption(sysin))";  
title2 " ";  
  
options nofmterr;  
  
*** Reading in the analysis datasets used for the DSIC;  
libname inlib5 xport  
"/prj/niddk/ims_analysis/NASH/private_orig_data/NASHCRN_Data_Sharing_PIVENS_NEJM2010/NASHCRN_Data  
_Sharing_PIVENS_NEJM2010/Datasets/table1.xpt";  
proc copy in =inlib5 out= work;  
libname inlib6 xport  
"/prj/niddk/ims_analysis/NASH/private_orig_data/NASHCRN_Data_Sharing_PIVENS_NEJM2010/NASHCRN_Data  
_Sharing_PIVENS_NEJM2010/Datasets/table2.xpt";  
proc copy in =inlib6 out= work;  
libname inlib7 xport  
"/prj/niddk/ims_analysis/NASH/private_orig_data/NASHCRN_Data_Sharing_PIVENS_NEJM2010/NASHCRN_Data  
_Sharing_PIVENS_NEJM2010/Datasets/table3.xpt";  
proc copy in =inlib7 out= work;  
  
*** Data from the Primary outcome paper that was converted to .csv format so that the DSIC data  
could be easily compared;  
FILENAME table1 '/prj/niddk/ims_analysis/NASH/private_created_data/nash_table1_data.csv';  
FILENAME table2 '/prj/niddk/ims_analysis/NASH/private_created_data/nash_table2_data.csv';  
FILENAME table3 '/prj/niddk/ims_analysis/NASH/private_created_data/nash_table3_data.csv';  
  
*** Output CSV files that will be converted to .xls before being added to the DSIC document;  
FILENAME out_t1 '/prj/niddk/ims_analysis/NASH/private_created_data/nash_table1_dsic.csv';  
FILENAME out_t2 '/prj/niddk/ims_analysis/NASH/private_created_data/nash_table2_dsic.csv';  
FILENAME out_t3 '/prj/niddk/ims_analysis/NASH/private_created_data/nash_table3_dsic.csv';  
  
%macro baseline_freq(dataset_name,var_name);  
  
    *** Creating a frequency table in the format of Table 1 in the primary outcome paper;  
    proc freq data = &dataset_name;  
        table (&var_name.)*TX;  
        title3 "Frequency table of the &var_name. variable in the analysis dataset";  
  
    *** Outputting the frequency data to work.&var_name._cross using the ODS output;  
    ods output CrossTabFreqs = work.&var_name._cross;  
  
    *** Creating two datasets (one per arm) so that the data can be in the correct format;  
    data &var_name._cross_Placebo &var_name._cross_VitE &var_name._cross_Piog  
&var_name._cross_all;  
        set &var_name._cross;  
        if TX = "Plbo" and &var_name. NE . then do; first_stat = round(colpercent,1);  
output &var_name._cross_Placebo; end;  
        else if TX = "VitE" and &var_name. NE . then do; first_stat = round(colpercent,1);  
output &var_name._cross_VitE; end;
```

```

        else if TX = "Pio" and &var_name. NE . then do; first_stat = round(colpercent,1);
output &var_name._cross_Piog; end;
        else if TX = "" and &var_name. NE . then do; first_stat = round(Percent,1);
output &var_name._cross_all; end;
        else if lengthn(TX) NE 0 and &var_name. NE . then abort;

*** Creating a dataset with the merged data with the variables that contain the order of the
statistics;
        data &var_name._merge;
            merge &var_name._cross_Placebo (in = in1 keep= table &var_name. first_stat rename =
(first_stat = first_stat_Placebo))
                &var_name._cross_VitE (in = in2 keep= table &var_name. first_stat rename =
(first_stat = first_stat_VitE ))
                &var_name._cross_Piog (in = in3 keep= table &var_name. first_stat rename =
(first_stat = first_stat_Piog ))
                &var_name._cross_all (in = in4 keep= table &var_name. first_stat rename =
(first_stat = first_stat_all ))
            ;
        by table &var_name.;
            if in1 and in2 and in3 and in4 then output &var_name._merge;
            else abort;

%mend;

%macro baseline_median(dataset_name,var_name);

        proc sort data = &dataset_name;
            by TX;

            *** Creating a means table in the format of Table 1 in the primary outcome paper that
            contain the median 25th percentile and 75th percentile;
            proc means data = &dataset_name mean Std ;
                var &var_name.;
                by TX;

                *** Outputting the statistics to the work.&var_name._summary dataset using the ODS
                output;
                ods output Summary = work.&var_name._means1;

                *** Creating two datasets (one per arm) so that the data can be in the correct format;
                data &var_name._means_Placebo &var_name._means_VitE &var_name._means_Piog;
                set &var_name._means1;
                length table_name $ 30.;
                table_name = upcase("&var_name.");
                if TX = "Plbo" then do; first_stat = round(&var_name._Mean,0.01); second_stat =
round(&var_name._StdDev,0.01);output &var_name._means_Placebo; end;
                else if TX = "VitE" then do; first_stat = round(&var_name._Mean,0.01); second_stat =
round(&var_name._StdDev,0.01);output &var_name._means_VitE; end;
                else if TX = "Pio" then do; first_stat = round(&var_name._Mean,0.01); second_stat =
round(&var_name._StdDev,0.01);output &var_name._means_Piog; end;

                proc means data = &dataset_name mean Std ;
                    var &var_name.;

                    *** Outputting the statistics to the work.&var_name._summary dataset using the ODS
                    output;
                    ods output Summary = work.&var_name._means_all;

                    data &var_name._means_all;
                    set &var_name._means_all;
                    length table_name $ 30.;
                    table_name = upcase("&var_name.");
                    first_stat = round(&var_name._Mean,0.01);
                    second_stat = round(&var_name._StdDev,0.01);

            *** Creating a dataset with the merged data with the variables that contain the order of the
            statistics;
            data &var_name._means;

```

```

merge &var_name._means_Placebo (in = in1 keep = table_name first_stat second_stat
rename = (first_stat = first_stat_Placebo second_stat = second_stat_Placebo))
&var_name._means_VitE (in = in2 keep = table_name first_stat second_stat
rename = (first_stat = first_stat_VitE second_stat = second_stat_VitE ))
&var_name._means_Piog (in = in2 keep = table_name first_stat second_stat
rename = (first_stat = first_stat_Piog second_stat = second_stat_Piog ))
&var_name._means_all (in = in2 keep = table_name first_stat second_stat
rename = (first_stat = first_stat_all second_stat = second_stat_all ))
;
by table_name;

%mend;

```

```

*****;
***** Check Table 1 *****;
*****;

```

```

*** Running the baseline_freq on the categorical variables in the Table 1 manuscript file;
%baseline_freq(table1,FEMALE );
%baseline_freq(table1,HISP );
%baseline_freq(table1,WHITE );
%baseline_freq(table1,BLBALL0 );

```

```

data compare_table1_freq(drop =table FEMALE HISP WHITE BLBALL0);
set FEMALE_merge (keep = table FEMALE first_stat_Placebo first_stat_VitE
first_stat_Piog first_stat_all where = (FEMALE = 1))
HISP_merge (keep = table HISP first_stat_Placebo first_stat_VitE
first_stat_Piog first_stat_all where = (HISP = 1))
WHITE_merge (keep = table WHITE first_stat_Placebo first_stat_VitE
first_stat_Piog first_stat_all where = (WHITE = 0))
BLBALL0_merge (keep = table BLBALL0 first_stat_Placebo first_stat_VitE
first_stat_Piog first_stat_all where = (BLBALL0 = 1))
;
length table_name $ 30.;
table_name = substr(table,7,length(table)-11);

```

```

*** Running the baseline_media on the continuous variables in the Table 1 manuscript file;

```

```

%baseline_median(table1,age );
%baseline_median(table1,PCS );
%baseline_median(table1,MCS );
%baseline_median(table1,ALT );
%baseline_median(table1,AST );
%baseline_median(table1,GGT );
%baseline_median(table1,ALKA );
%baseline_median(table1,BILIT );
%baseline_median(table1,TRIG );
%baseline_median(table1,CHOL );
%baseline_median(table1,HDL );
%baseline_median(table1,LDL );
%baseline_median(table1,GLUC );
%baseline_median(table1,HOMA );
%baseline_median(table1,WEIGHT );
%baseline_median(table1,BMI );
%baseline_median(table1,WAISTCM );
%baseline_median(table1,TOTF );
%baseline_median(table1,BLNAS );
%baseline_median(table1,BLSTEATO );
%baseline_median(table1,BLINFLAM );
%baseline_median(table1,BLBALL );
%baseline_median(table1,BLFIBRO );

```

```

data compare_table1_mean;
set age_means
PCS_means
MCS_means
ALT_means

```

```

    AST_means
    GGT_means
    ALKA_means
    BILIT_means
    TRIG_means
    CHOL_means
    HDL_means
    LDL_means
    GLUC_means
    HOMA_means
    WEIGHT_means
    BMI_means
    WAISTCM_means
    TOTF_means
    BLNAS_means
    BLSTEATO_means
    BLINFLAM_means
    BLBALL_means
    BLFIBRO_means

;

data compare_table1;
  set compare_table1_freq compare_table1_mean;
  length char_compare_stat1-char_compare_stat4 $ 30.;
  if table_name in ('FEMALE' 'HISP' 'WHITE' 'BLBALL0' ) then do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.)) ;
    char_compare_stat2 = strip(put(first_stat_VitE ,8.)) ;
    char_compare_stat3 = strip(put(first_stat_Piog ,8.)) ;
    char_compare_stat4 = strip(put(first_stat_all ,8.)) ;
  end;
  else if table_name in ('BILIT') then do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.2)) || " , " ||
strip(put(second_stat_Placebo,8.2));
    char_compare_stat2 = strip(put(first_stat_VitE ,8.2)) || " , " ||
strip(put(second_stat_VitE ,8.2));
    char_compare_stat3 = strip(put(first_stat_Piog ,8.2)) || " , " ||
strip(put(second_stat_Piog ,8.2));
    char_compare_stat4 = strip(put(first_stat_all ,8.2)) || " , " ||
strip(put(second_stat_all ,8.2));
  end;
  else if table_name in ('AGE' 'HOMA' 'BLNAS' 'BLSTEATO' 'BLINFLAM' 'BLBALL' 'BLFIBRO') then do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.1)) || " , " ||
strip(put(second_stat_Placebo,8.1));
    char_compare_stat2 = strip(put(first_stat_VitE ,8.1)) || " , " ||
strip(put(second_stat_VitE ,8.1));
    char_compare_stat3 = strip(put(first_stat_Piog ,8.1)) || " , " ||
strip(put(second_stat_Piog ,8.1));
    char_compare_stat4 = strip(put(first_stat_all ,8.1)) || " , " ||
strip(put(second_stat_all ,8.1));
  end;
  else do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.)) || " , " ||
strip(put(second_stat_Placebo,8.));
    char_compare_stat2 = strip(put(first_stat_VitE ,8.)) || " , " ||
strip(put(second_stat_VitE ,8.));
    char_compare_stat3 = strip(put(first_stat_Piog ,8.)) || " , " ||
strip(put(second_stat_Piog ,8.));
    char_compare_stat4 = strip(put(first_stat_all ,8.)) || " , " ||
strip(put(second_stat_all ,8.));
  end;

*** Importing the Table 1 Data taken from the primary outcome paper;
data table1_data;
  infile table1 delimiter = ',' MISSOVER DSD firstobs=3 ls=1080;
  length characteristic $45 table_name $30 char_stat1-char_stat4 $23 ;
  input
  characteristic $ table_name $ char_stat1 $ char_stat2 $ char_stat3 $ char_stat4 $ stat1
  stat2 stat3 stat4 stat5 stat6 stat7 stat8 stat9 $

```



```

;
  if lengthn(characteristic) NE 0 then output table1_data;

data table1_data;
  set table1_data;
  sort_order = _n_;
  drop stat9;
  table_name = upcase(table_name);
  char_stat1 = translate(char_stat1, ' ', '?');
  char_stat2 = translate(char_stat2, ' ', '?');
  char_stat3 = translate(char_stat3, ' ', '?');
  char_stat4 = translate(char_stat4, ' ', '?');

*** Merging the DSIC Table 1 data and the Table 1 data from the manuscript;
*** Creating variables to calculate the difference between the datasets;
proc sort data = compare_table1;
  by table_name ;
proc sort data = table1_data;
  by table_name ;

data combined_table1_dataset;
  merge compare_table1 (in = in1)
        table1_data    (in = in2)
        ;
  by table_name ;
  if in1 and in2 then output combined_table1_dataset;
  else abort;

data combined_table1_dataset_0 combined_table1_dataset_1 combined_table1_dataset_2;
  set combined_table1_dataset;
  if table_name in ('BILIT') then output combined_table1_dataset_2;
  else if table_name in ('AGE' 'HOMA' 'BLNAS' 'BLSTEATO' 'BLINFLAM' 'BLBALL' 'BLFIBRO' ) then
output combined_table1_dataset_1;
  else output combined_table1_dataset_0;

data combined_table1_dataset_0;
  set combined_table1_dataset_0;
  length char_diff1-char_diff4 $ 30.;
  if table_name in ('FEMALE' 'HISP' 'WHITE' 'BLBALL0') then do;
    diff1 = round((round(stat1,1) - round(first_stat_Placebo ,1)),1);
    diff3 = round((round(stat3,1) - round(first_stat_VitE ,1)),1);
    diff5 = round((round(stat5,1) - round(first_stat_Piog ,1)),1);
    diff7 = round((round(stat7,1) - round(first_stat_all ,1)),1);
    char_diff1 = strip(put(diff1,8.)) ;
    char_diff2 = strip(put(diff3,8.)) ;
    char_diff3 = strip(put(diff5,8.)) ;
    char_diff4 = strip(put(diff7,8.)) ;
  end;
  else do;
    diff1 = round((round(stat1,1) - round(first_stat_Placebo ,1)),1);
    diff2 = round((round(stat2,1) - round(second_stat_Placebo,1)),1);
    diff3 = round((round(stat3,1) - round(first_stat_VitE ,1)),1);
    diff4 = round((round(stat4,1) - round(second_stat_VitE ,1)),1);
    diff5 = round((round(stat5,1) - round(first_stat_Piog ,1)),1);
    diff6 = round((round(stat6,1) - round(second_stat_Piog ,1)),1);
    diff7 = round((round(stat7,1) - round(first_stat_all ,1)),1);
    diff8 = round((round(stat8,1) - round(second_stat_all ,1)),1);
    char_diff1 = strip(put(diff1,8.)) || " , " || strip(put(diff2,8.));
    char_diff2 = strip(put(diff3,8.)) || " , " || strip(put(diff4,8.));
    char_diff3 = strip(put(diff5,8.)) || " , " || strip(put(diff6,8.));
    char_diff4 = strip(put(diff7,8.)) || " , " || strip(put(diff8,8.));
  end;

data combined_table1_dataset_1;
  set combined_table1_dataset_1;
  length char_diff1-char_diff4 $ 30.;
  diff1 = round((round(stat1,0.1) - round(first_stat_Placebo ,0.1)),0.1);
  if stat2 ne . then diff2 = round((round(stat2,0.1) - round(second_stat_Placebo,0.1)),0.1);

```

```

        diff3 = round((round(stat3,0.1) - round(first_stat_VitE ,0.1)),0.1);
if stat4 ne . then diff4 = round((round(stat4,0.1) - round(second_stat_VitE ,0.1)),0.1);
        diff5 = round((round(stat5,0.1) - round(first_stat_Piog ,0.1)),0.1);
if stat6 ne . then diff6 = round((round(stat6,0.1) - round(second_stat_Piog ,0.1)),0.1);
        diff7 = round((round(stat7,0.1) - round(first_stat_all ,0.1)),0.1);
if stat8 ne . then diff8 = round((round(stat8,0.1) - round(second_stat_all ,0.1)),0.1);
        char_diff1 = strip(put(diff1,8.1)) || " , " || strip(put(diff2,8.1));
        char_diff2 = strip(put(diff3,8.1)) || " , " || strip(put(diff4,8.1));
        char_diff3 = strip(put(diff5,8.1)) || " , " || strip(put(diff6,8.1));
        char_diff4 = strip(put(diff7,8.1)) || " , " || strip(put(diff8,8.1));

data combined_table1_dataset_2;
  set combined_table1_dataset_2;
  length char_diff1-char_diff4 $ 30.;
        diff1 = round((round(stat1,0.01) -
round(first_stat_Placebo ,0.01)),0.01);
  if stat2 ne . then diff2 = round((round(stat2,0.01) -
round(second_stat_Placebo,0.01)),0.01);
        diff3 = round((round(stat3,0.01) -
round(first_stat_VitE ,0.01)),0.01);
  if stat4 ne . then diff4 = round((round(stat4,0.01) -
round(second_stat_VitE ,0.01)),0.01);
        diff5 = round((round(stat5,0.01) -
round(first_stat_Piog ,0.01)),0.01);
  if stat6 ne . then diff6 = round((round(stat6,0.01) -
round(second_stat_Piog ,0.01)),0.01);
        diff7 = round((round(stat7,0.01) -
round(first_stat_all ,0.01)),0.01);
  if stat8 ne . then diff8 = round((round(stat8,0.01) -
round(second_stat_all ,0.01)),0.01);

        char_diff1 = strip(put(diff1,8.2)) || " , " || strip(put(diff2,8.2));
        char_diff2 = strip(put(diff3,8.2)) || " , " || strip(put(diff4,8.2));
        char_diff3 = strip(put(diff5,8.2)) || " , " || strip(put(diff6,8.2));
        char_diff4 = strip(put(diff7,8.2)) || " , " || strip(put(diff8,8.2));

data combined_table1_dataset_out;
  set combined_table1_dataset_0 combined_table1_dataset_1 combined_table1_dataset_2;
  label
  char_stat1 = "Placebo (N=83) [Manuscript]"
  char_compare_stat1 = "Placebo (N=83) [DSIC] "
  char_diff1 = "Placebo (N=83) [Difference]"
  char_stat2 = "Vitamin E (N=84) [Manuscript]"
  char_compare_stat2 = "Vitamin E (N=84) [DSIC] "
  char_diff2 = "Vitamin E (N=84) [Difference]"
  char_stat3 = "Pioglitazone (N=80) [Manuscript]"
  char_compare_stat3 = "Pioglitazone (N=80) [DSIC] "
  char_diff3 = "Pioglitazone (N=80) [Difference]"
  char_stat4 = "Total (N=247) [Manuscript]"
  char_compare_stat4 = "Total (N=247) [DSIC] "
  char_diff4 = "Total (N=247) [Difference]"
;

*****
***** Check Table 2 *****
*****

%baseline_freq(table2, IMPROVE );
%baseline_freq(table2, ISTEATO );
%baseline_freq(table2, IINFLAM );
%baseline_freq(table2, IBALL );
%baseline_freq(table2, IFIBRO );
%baseline_freq(table2, RESOLVE );

proc freq data = table2;
  tables tx * (IMPROVE RESOLVE) ;

```

```

    *** Outputting the frequency data to work.&var_name._cross using the ODS output;
ods output CrossTabFreqs = work.RESOLVE_cross;

data RESOLVE_cross (keep = table TX Frequency table_name) ;
  set RESOLVE_cross;
  if ((index(table,'IMPROVE') > 0 and IMPROVE = .) or (index(table,'RESOLVE') > 0 and RESOLVE
= .)) and TX ne '';
  length table_name $ 30.;
  table_name = substr(table,12,length(table)-11);
  if table_name = 'IMPROVE' then table_name = 'CASE_COUNT';
  else if table_name = 'RESOLVE' then table_name = 'RESOLVE_COUNT';

*proc print data = RESOLVE_cross;
* title3 'RESOLVE_cross';

proc transpose data=RESOLVE_cross out=RESOLVE_cross_trans ;
  by table_name;

*proc print data = RESOLVE_cross_trans;
* title3 'RESOLVE_cross_trans';

data compare_table2_freq(drop =table IMPROVE ISTEATO IINFLAM IBALL IFIBRO RESOLVE);
  set IMPROVE_merge (keep = table IMPROVE first_stat_Placebo first_stat_VitE
first_stat_Piog where = (IMPROVE = 1))
  ISTEATO_merge (keep = table ISTEATO first_stat_Placebo first_stat_VitE
first_stat_Piog where = (ISTEATO = 1))
  IINFLAM_merge (keep = table IINFLAM first_stat_Placebo first_stat_VitE
first_stat_Piog where = (IINFLAM = 1))
  IBALL_merge (keep = table IBALL first_stat_Placebo first_stat_VitE
first_stat_Piog where = (IBALL = 1))
  IFIBRO_merge (keep = table IFIBRO first_stat_Placebo first_stat_VitE
first_stat_Piog where = (IFIBRO = 1))
  RESOLVE_merge (keep = table RESOLVE first_stat_Placebo first_stat_VitE
first_stat_Piog where = (RESOLVE = 1))
  RESOLVE_cross_trans (in = in1 keep = table_name coll-col3 rename = (COL1 =first_stat_Piog
COL2 =first_stat_Placebo COL3=first_stat_VitE))
;
  length table_name $ 30.;
  if not in1 then table_name = substr(table,7,length(table)-11);

*proc print data = compare_table2_freq;
* title3 'compare_table2_freq';

proc sort data = table2;
  by tx;

proc means data = table2 mean Std ;
  var CSTEATO CINFLAM CBALL CNAS CFIBRO;
  by TX;

%baseline_median(table2,CSTEATO);
%baseline_median(table2,CBALL );
%baseline_median(table2,CINFLAM);
%baseline_median(table2,CNAS );
%baseline_median(table2,CFIBRO );

data compare_table2_mean(keep = table_name first_stat_Placebo first_stat_VitE first_stat_Piog );
  set CSTEATO_means
  CINFLAM_means
  CBALL_means
  CNAS_means
  CFIBRO_means
;

data compare_table2;
  set compare_table2_freq compare_table2_mean;
  length char_compare_stat1-char_compare_stat3 $ 30.;

```

```

if table_name in ('CSTEATO' 'CINFLAM' 'CBALL' 'CNAS' 'CFIBRO') then do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.1)) ;
    char_compare_stat2 = strip(put(first_stat_VitE ,8.1)) ;
    char_compare_stat3 = strip(put(first_stat_Piog ,8.1)) ;
end;
else do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.)) ;
    char_compare_stat2 = strip(put(first_stat_VitE ,8.)) ;
    char_compare_stat3 = strip(put(first_stat_Piog ,8.)) ;
end;

*** Importing the Table 2 Data taken from the primary outcome paper;
data table2_data;
    infile table2 delimiter = ',' MISSOVER DSD firstobs=1 ls=1080;
    length characteristic $45 table_name $30 char_stat1-char_stat3 char_statp1-char_statp2 $23 ;
    input
    characteristic $ table_name $ char_stat1 $ char_stat2 $ char_stat3 $ stat1 stat2 stat3
    char_statp1 $ char_statp2 $
;
    if lengthn(characteristic) NE 0 then output table2_data;

data table2_data;
    set table2_data;
    sort_order = _n_;

*** Merging the DSIC Table 1 data and the Table 1 data from the manuscript;
*** Creating variables to calculate the difference between the datasets;
proc sort data = compare_table2;
    by table_name ;
proc sort data = table2_data;
    by table_name ;

data combined_table2_dataset;
    merge compare_table2 (in = in1)
          table2_data (in = in2)
          ;
    by table_name ;
    if in1 and in2 then output combined_table2_dataset;
    else abort;

data combined_table2_dataset;
    set combined_table2_dataset;
    if table in ('IMPROVE' 'ISTEATO' 'IINFLAM' 'IBALL' 'IFIBRO' 'RESOLVE' 'CASE_COUNT'
'RESOLVE_COUNT') then do;
        diff1 = round((round(stat1,1) - round(first_stat_Placebo ,1)),1);
        diff2 = round((round(stat2,1) - round(first_stat_VitE ,1)),1);
        diff3 = round((round(stat3,1) - round(first_stat_Piog ,1)),1);
        char_diff1 = strip(put(diff1,8.)) ;
        char_diff2 = strip(put(diff2,8.)) ;
        char_diff3 = strip(put(diff3,8.)) ;
    end;
    else do;
        diff1 = round((round(stat1,0.1) - round(first_stat_Placebo ,0.1)),0.1);
        diff2 = round((round(stat2,0.1) - round(first_stat_VitE ,0.1)),0.1);
        diff3 = round((round(stat3,0.1) - round(first_stat_Piog ,0.1)),0.1);
        char_diff1 = strip(put(diff1,8.1)) ;
        char_diff2 = strip(put(diff2,8.1)) ;
        char_diff3 = strip(put(diff3,8.1)) ;
    end;
label
    char_stat1 = "Placebo [Manuscript]"
    char_compare_stat1 = "Placebo [DSIC]"
    char_diff1 = "Placebo [Difference]"
    char_stat2 = "Vitamin E [Manuscript]"
    char_compare_stat2 = "Vitamin E [DSIC]"
    char_diff2 = "Vitamin E [Difference]"
    char_stat3 = "Pioglitazone [Manuscript]"
    char_compare_stat3 = "Pioglitazone [DSIC]"
    char_diff3 = "Pioglitazone [Difference]"

```

```

;

*****
***** Check Table 3 *****
*****

proc sort data = table3;
  by tx;
  where not missing(alt);

proc means data = table3 mean Std ;
  var CALT CAST CGGT CALKA CBILIT  CTRIG CCHOL CHDL CLDL CGLUC CHOMA CWEIGHT CBMI CWAIST CTOTF
  CPCS CMCS ;
  by TX;

%baseline_median(table3,CALT          );
%baseline_median(table3,CAST          );
%baseline_median(table3,CGGT          );
%baseline_median(table3,CALKA         );
%baseline_median(table3,CBILIT        );
%baseline_median(table3,CTRIG         );
%baseline_median(table3,CCHOL         );
%baseline_median(table3,CHDL          );
%baseline_median(table3,CLDL          );
%baseline_median(table3,CGLUC         );
%baseline_median(table3,CHOMA         );
%baseline_median(table3,CWEIGHT       );
%baseline_median(table3,CBMI          );
%baseline_median(table3,CWAIST        );
%baseline_median(table3,CTOTF         );
%baseline_median(table3,CPCS          );
%baseline_median(table3,CMCS          );

data compare_table3(keep = table_name first_stat_Placebo first_stat_VitE first_stat_Piog
char_compare_stat1-char_compare_stat3);
  set CALT_means
      CAST_means
      CGGT_means
      CALKA_means
      CBILIT_means
      CTRIG_means
      CCHOL_means
      CHDL_means
      CLDL_means
      CGLUC_means
      CHOMA_means
      CWEIGHT_means
      CBMI_means
      CWAIST_means
      CTOTF_means
      CPCS_means
      CMCS_means
;
  length char_compare_stat1-char_compare_stat3 $ 30.;
  if table_name = 'CBILIT' then do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.2)) ;
    char_compare_stat2 = strip(put(first_stat_VitE      ,8.2)) ;
    char_compare_stat3 = strip(put(first_stat_Piog      ,8.2)) ;
  end;
  else do;
    char_compare_stat1 = strip(put(first_stat_Placebo,8.1)) ;
    char_compare_stat2 = strip(put(first_stat_VitE      ,8.1)) ;
    char_compare_stat3 = strip(put(first_stat_Piog      ,8.1)) ;
  end;

proc print data = compare_table3;
  title3 'compare_table3';

```

```

*** Importing the Table 3 Data taken from the primary outcome paper;
data table3_data;
  infile table3 delimiter = ',' MISSOVER DSD firstobs=1 ls=1080;
  length characteristic $45 table_name $30 char_stat1-char_stat3 char_statp1-char_statp2 $23 ;
  input
  characteristic $ table_name $ char_stat1 $ char_stat2 $ char_stat3 $ stat1 stat2 stat3
  char_statp1 $ char_statp2 $
;
  if lengthn(characteristic) NE 0 then output table3_data;

data table3_data;
  set table3_data;
  sort_order = _n_;

proc print data = table3_data;
  title3 'table3_data';

*** Merging the DSIC Table 3 data and the Table 1 data from the manuscript;
*** Creating variables to calculate the difference between the datasets;
proc sort data = compare_table3;
  by table_name ;
proc sort data = table3_data;
  by table_name ;

data combined_table3_dataset;
  merge compare_table3 (in = in1)
        table3_data (in = in2)
  ;
  by table_name ;
  if in1 and in2 then output combined_table3_dataset;
  else abort;

data combined_table3_dataset;
  set combined_table3_dataset;
  if table in ('CBILIT') then do;
    diff1 = round((round(stat1,0.01) - round(first_stat_Placebo ,0.01)),0.01);
    diff2 = round((round(stat2,0.01) - round(first_stat_VitE ,0.01)),0.01);
    diff3 = round((round(stat3,0.01) - round(first_stat_Piog ,0.01)),0.01);
    char_diff1 = strip(put(diff1,8.2)) ;
    char_diff2 = strip(put(diff2,8.2)) ;
    char_diff3 = strip(put(diff3,8.2)) ;
  end;
  else do;
    diff1 = round((round(stat1,0.1) - round(first_stat_Placebo ,0.1)),0.1);
    diff2 = round((round(stat2,0.1) - round(first_stat_VitE ,0.1)),0.1);
    diff3 = round((round(stat3,0.1) - round(first_stat_Piog ,0.1)),0.1);
    char_diff1 = strip(put(diff1,8.1)) ;
    char_diff2 = strip(put(diff2,8.1)) ;
    char_diff3 = strip(put(diff3,8.1)) ;
  end;
  label
  char_stat1 = "Placebo [Manuscript]"
  char_compare_stat1 = "Placebo [DSIC] "
  char_diff1 = "Placebo [Difference]"
  char_stat2 = "Vitamin E [Manuscript]"
  char_compare_stat2 = "Vitamin E [DSIC] "
  char_diff2 = "Vitamin E [Difference]"
  char_stat3 = "Pioglitazone [Manuscript]"
  char_compare_stat3 = "Pioglitazone [DSIC] "
  char_diff3 = "Pioglitazone [Difference]"
;
*** Outputting the data to a csv format to be added to the DSIC;

proc sort data = combined_table1_dataset_out ;
  by sort_order;

ods csv file = out_t1;
run;

```

```

proc print data = combined_table1_dataset_out NOOBS label;
    var characteristic
    char_stat1
    char_compare_stat1
    char_diff1
    char_stat2
    char_compare_stat2
    char_diff2
    char_stat3
    char_compare_stat3
    char_diff3
    char_stat4
    char_compare_stat4
    char_diff4
;
    title3 "DSIC Check of Table 1: Baseline Characteristic of the Study Subjects";
run;

proc sort data = combined_table2_dataset ;
    by sort_order;

ods csv file = out_t2;
run;

proc print data = combined_table2_dataset NOOBS label;
    var characteristic
    char_stat1
    char_compare_stat1
    char_diff1
    char_stat2
    char_compare_stat2
    char_diff2
    char_stat3
    char_compare_stat3
    char_diff3
;
    title3 "DSIC Check of Table 2: Primary Outcome and CHanges in Histologic Features of the
Liver after 96 Weeks of Treatment.";
run;

proc sort data = combined_table3_dataset ;
    by sort_order;

ods csv file = out_t3;
run;

proc print data = combined_table3_dataset NOOBS label;
    var characteristic
    char_stat1
    char_compare_stat1
    char_diff1
    char_stat2
    char_compare_stat2
    char_diff2
    char_stat3
    char_compare_stat3
    char_diff3
;
    title3 "DSIC Check of Table 3: Primary Outcome and CHanges in Histologic Features of the
Liver after 96 Weeks of Treatment.";
run;

```