Data Set Integrity Check for the Trial of Botulinum Toxin Injection for the Management of BPH (MIST2) Data Files

Prepared by Jane Wang
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton MD 20705
April 29, 2014

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

Benign prostatic hyperplasia (BPH) and lower urinary tract symptoms are bothersome and costly conditions affecting nearly half of men older than age 50 years and 90% of men older than age 80. In the search for more effective noninvasive therapies, the role of the neural activity in BPH has been studied. Botulinum toxin injection has been used for several urological conditions including voiding dysfunction, and studies have shown that intraprostatic injection may be beneficial in lower urinary tract symptoms secondary to BPH.

Eligible men were randomized in a 1:1 ratio to a 100 or a 300 U dose of onabotulinum toxin A. The study was double-blind as to the toxin dose. The drug was injected transrectally into the prostatic peripheral transitional zone with ultrasound guidance. It was administered as 2 cc per lobe in 1 ml aliquots at 2 injection sites for a total of 4 sites.

Men were contacted by telephone at 1 week after treatment, and evaluated in person at 1, 2, 3, 6, 9 and 12 months. Uroflowmetry, vital signs, AUASI and medical conditions were collected or performed at baseline and each visit; transrectal ultrasound and questionnaires on the impact of BPH, bladder function, erectile function and ejaculatory function were administered at baseline, and 3 and 12 months; bladder ultrasound was performed at baseline and 12 months; and physical and digital rectal examination, and prostate specific antigen were determined at baseline, and at 3, 6 and 12 months.

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the MIST "Data" folder included in the data package. For this replication, variables were taken from following data sets:

1. mist2_aua_052110
2. mist2_tx0_052110
3. mist2_scr1_052110
4. mist2_m06_052110
5. mist2_w4w8m9_052110
6. mist2_w12m12_052110l

4 Statistical Methods

Analyses were performed to duplicate results from the data published by **Crawford** et al [1] J Urol. 2011 September ; 186(3): 965–970. doi:10.1016/j.juro.2011.04.062.

To verify the integrity of the MIST datasets, descriptive statistics of baseline characteristics and 12 month follow-up visit were computed, by treatment group.

5 Results

Table 1 in the publication [1] AUASI and Qmax at baseline and follow up. 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 2. The results of the replication are similar to the published results.

Table 2 in the publication [1] Baseline clinical and demographic characteristics. 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 similar to the published results.

Table 3 in the publication [1]: Baseline comparison of completers vs dropouts. 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 the published results.

Note that the adverse event data that was used to create the adverse event table in the publication was not included in this data package.

6 Conclusions

The NIDDK repository is confident that the MIST2 data files to be distributed are a true copy of the study data.

7 References

Effects of 100 and 300 Units of Onabotulinum Toxin A on Lower Urinary Tract Symptoms of Benign Prostatic Hyperplasia: A Phase II Randomized Clinical Trial. *J Urol. 2011 September ; 186(3): 965–970. doi:10.1016/j.juro.2011.04.062.*[1]

Table A: Variables used to replicate Table 1: AUASI and Qmax at baseline and follow up

Variable for AUASI are from data set MIST2_AUA_052110
aunempty
autwohrs
aupostpo
austopst
auweakst
aupushst
augetupn

Table 1 Variable for Qmax	**Variables Used in Replication**	**Data set variable from**
BASELINE	s1maxrt	mist2_scr1_052110
M01	w4maxflo	mist2_w4w8m9_052110
M02	w4maxflo	mist2_w4w8m9_052110
M03	wmmaxflo	mist2_w12m12_052110
M06	m6maxflo	mist2_m06_052110
M09	w4maxflo	mist2_w4w8m9_052110
M12	wmmaxflo	mist2_w12m12_052110

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

	AUASI 100U [Manuscript]	AUASI 100U [DSIC]	AUASI 100U [Difference]	AUASI 300U [Manuscript]	AUASI 300U [DSIC]	AUASI 300U [Difference]
BASELINE	18.8	19.2	-0.4	19.5	20.4	-0.9
M01	12	13.4	-1.4	12.5	13.1	-0.6
M02	11.1	12.7	-1.6	12.4	12.9	-0.5
M03	11.7	12.9	-1.2	10.6	11.2	-0.6
M06	12	12.6	-0.6	12.3	12.4	-0.1
M09	12.2	12.4	-0.2	13.2	13.9	-0.7
M12	11.9	12	-0.1	12.4	12.2	0.2

	Qmax 100U [Manuscript]	Qmax 100U [DSIC]	Qmax 100U [Difference]	Qmax 300U [Manuscript]	Qmax 300U [DSIC]	Qmax 300U [Difference]
BASELINE	10	10.2	-0.2	9.6	10.1	-0.5
M01	12.6	11.9	0.7	12.1	11.7	0.4
M02	11.3	11	0.3	12.8	12.2	0.6
M03	12.5	12.1	0.4	12.2	12.2	0
M06	12.4	12	0.4	12.9	12.3	0.6
M09	13	12.7	0.3	12.1	11.7	0.4
M12	12.2	11.6	0.6	11.9	11.6	0.3

Table C: Variables used to replicate Table 2: Baseline clinical and demographic characteristics

Table Variable	**Variables Used in Replication**	**Data set variable from**
age	age	MIST2_SCR1_052110
Race/ethnicity white nonHispanic	raceeth	MIST2_SCR1_052110
Education beyond high school		
married or committed relationship	s1marrie	MIST2_SCR1_052110
yrs with BPH symptoms	s1yrsbph	mist2_scr1_052110
AUASI	aunempty autwohrs aupostpo austopst auweakst aupushst augetupn	MIST2_AUA_052110
ml/sec Qmax	s1maxrt	mist2_scr1_052110
mm prostate length	txlength	mist2_tx0_052110
cc prostate vol	txprovol	mist2_tx0_052110
cm2 max sagittal area	txmaxima	mist2_tx0_052110

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

	100U [Manuscript]	**100U [DSIC]**	**100U [Difference]**	**300U [Manuscript]**	**300U [DSIC]**	**300U [Difference]**
GROUP	.	100	.	.	300	.
Mean age	67	67	0	65.3	65.3	0
SD age	8.6	8.6	0	7.7	7.7	0
% Race/ethnicity white nonHispanic	92.6	92.6	0	86.2	87.9	-1.7
% Education beyond high school	86.7	.	.	87.7	.	.
% In married or committed relationship	77.9	77.9	0	78.5	78.5	0
Mean yrs with BPH symptoms	7.9	7.9	0	8.3	8.3	0
SD yrs with BPH symptoms	6.1	6.1	0	6.3	6.3	0
Mean AUASI	19.2	19.2	0	20.1	20.4	-0.3
SD AUASI	5.9	5.8	0.1	6.5	6.7	-0.2
Mean ml/sec Qmax	10	10.2	-0.2	9.7	10.1	-0.4
SD ml/sec Qmax	3	2.9	0.1	2.5	2.9	-0.4
Mean mm prostate length	44.4	44.4	0	45	45	0
SD mm prostate length	15.4	15.4	0	15.1	15.1	0
Mean cc prostate vol	51.1	51.1	0	48.1	48.1	0
SD cc prostate vol	24.2	24.2	0	26.7	26.7	0
Mean cm2 max sagittal area	15.3	15.3	0	14	14	0
SD cm2 max sagittal area	5.8	5.8	0	6.3	6.3	0

Table E: Variables used to replicate Table 3 Baseline comparison of completers vs dropouts

Table Variable	**Variables Used in Replication**	**Data set variable from**
age	age	MIST2_SCR1_052110
AUASI	aunempty autwohrs aupostpo austopst auweakst aupushst augetupn	MIST2_AUA_052110
mm prostate length	txlength	mist2_tx0_052110
cc prostate vol	txprovol	mist2_tx0_052110

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

	12-mo completers (108) [Manuscript]	**12-mo completers (108) [DSIC]**	**12-mo completers (108) [Difference]**	**Dropouts(26) [Manuscript]**	**Dropouts(26) [DSIC]**	**Dropouts(26) [Difference]**
age	66.0 (8.2)	66.0 (8.2)	0.0 (0.0)	67.0 (8.3)	67.0 (8.3)	0.0 (0.0)
AUASI	19.2 (6.0)	19.1 (6.1)	0.1 (-0.1)	21.5 (6.8)	22.5 (6.5)	-1.0 (0.3)
mm prostate length	9.8 (2.7)	10.4 (2.8)	-0.6 (-0.1)	10.1 (3.1)	9.3 (3.4)	0.8 (-0.3)
cc prostate vol	50.1 (26.7)	50.1 (26.7)	0.0 (0.0)	47.7 (19.8)	47.7 (19.8)	0.0 (0.0)

Attachment A: SAS Code

```
/*************************************************************************************************
**********************
***Program: /prj/niddk/ims_analysis/MIST/prog_initial_analysis/MIST_integrity_check_table4.sas;
***Programmer: Jane Wang
***Date Created: 2/6/2014
***Purpose: To perform a Dataset Integrity Check (DSIC) between the MIST data and the primary
outcome paper:
***                          Effects of 100 and 300 Units of Onabotulinum Toxin A on Lower
Urinary Tract Symptoms of Benign Prostatic Hyperplasia:
***         A Phase II Randomized Clinical Trial
***         E. David Crawford*, Kathryn Hirst?, John W. Kusek, Robert F. Donnell?, Steven A.
Kaplan§, Kevin T. McVary||,
***         Lance A. Mynderse?, Claus G. Roehrborn**, Christopher P. Smith??, and Reginald
Bruskewitz
***        The numbers in Tablesof the primary outcome paper will compared to the MIST data
received;
*************************************************************************************************
**********************/;

title1 "%sysfunc(getoption(sysin))";
title2 " ";

options nofmterr linesize=180;

*** Location of the MIST SAS dataset;

options nofmterr;

*** Reading in the analysis datasets used for the DSIC;
libname sas_data "/prj/niddk/ims_analysis/MIST/private_orig_data/MIST/SAS_Data_Files";
data mist2_aua_052110       ; set sas_data.mist2_aua_052110    ;
data mist2_tx0_052110       ; set sas_data.mist2_tx0_052110    ;
data mist2_scr1_052110      ; set sas_data.mist2_scr1_052110   ;
data mist2_m06_052110       ; set sas_data.mist2_m06_052110    ;
data mist2_w4w8m9_052110    ; set sas_data.mist2_w4w8m9_052110 ;
data mist2_w12m12_052110    ; set sas_data.mist2_w12m12_052110 ;

*** 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/MIST/private_created_data/mist_table1.csv';
FILENAME table2  '/prj/niddk/ims_analysis/MIST/private_created_data/mist_table2.csv';
FILENAME table3  '/prj/niddk/ims_analysis/MIST/private_created_data/mist_table3.csv';

*** Output CSV files that will be converted to .xls before being added to the DSIC document;
FILENAME out_t1  '/prj/niddk/ims_analysis/MIST/private_created_data/mist_table1_dsic.csv';
FILENAME out_t2  '/prj/niddk/ims_analysis/MIST/private_created_data/mist_table2_dsic.csv';
FILENAME out_t3  '/prj/niddk/ims_analysis/MIST/private_created_data/mist_table3_dsic.csv';


data mist2_aua_052110(keep = releaseid visitid total_aua);
  set mist2_aua_052110;
  total_aua = sum(aunempty, autwohrs, aupostpo, austopst, auweakst, aupushst, augetupn);

data grou_id (keep = group releaseid );
  set mist2_tx0_052110;
/*
proc sort data = grou_id;
  by releaseid;
proc sort data = mist2_aua_052110;
  by releaseid;

data mist2_aua_052110_group;
  merge mist2_aua_052110  grou_id;
  by releaseid;
*/
data
aua_M06(rename = (total_aua = M06_aua))
aua_M09(rename = (total_aua = M09_aua))
aua_M12(rename = (total_aua = M12_aua))
```

```
aua_S01(rename = (total_aua = baseline_aua))
aua_S02(rename = (total_aua = baseline_aua))
aua_M01(rename = (total_aua = M01_aua))
aua_M02(rename = (total_aua = M02_aua))
aua_M03(rename = (total_aua = M03_aua))
;
  set mist2_aua_052110;
       if visitid = 'M06' then output aua_M06 ;
  else if visitid = 'M09' then output aua_M09 ;
  else if visitid = 'M12' then output aua_M12 ;
  else if visitid = 'S01' then output aua_S01 ;
  else if visitid = 'S02' then output aua_S02 ;
  else if visitid = 'W04' then output aua_M01 ;
  else if visitid = 'W08' then output aua_M02 ;
  else if visitid = 'W12' then output aua_M03 ;

proc sort data = mist2_tx0_052110(keep = group releaseid txlength txprovol txmaxima rename =
(txlength = p_len_baseline txprovol = p_vol_baselin txmaxima = msa_baseline)) nodupkey;
  by releaseid;

proc sort data = mist2_aua_052110;
  by releaseid;

data aua prob;
  merge mist2_aua_052110 (in = in1) mist2_tx0_052110 (in = in2 );
  by releaseid;
  if in1 and in2 then output aua;
  else output prob;

proc freq data = aua;
  tables group * visitid/list missing;

data baseline_other (keep =  releaseid age raceeth s1marrie s1yrsbph s1maxrt rename = (s1yrsbph =
bph_baseline s1maxrt = baseline_qmax));
  set mist2_scr1_052110;




data M03 (keep = releaseid wmmaxflo rename = (wmmaxflo=M03_qmax))
     M12 (keep = releaseid wmmaxflo rename = (wmmaxflo=m12_qmax))  ;
  set mist2_w12m12_052110;
  if visitid = 'W12' then output m03;
  else if visitid = 'M12' then output m12;
  else abort;

data m06(keep = releaseid M06_qmax ) ;
  set mist2_m06_052110;
  M06_qmax =  m6maxflo;



data m01 (keep = releaseid w4maxflo rename = (w4maxflo = m01_qmax))
     m02 (keep = releaseid w4maxflo rename = (w4maxflo = m02_qmax))
     m09 (keep = releaseid w4maxflo rename = (w4maxflo = m09_qmax)) ;
  set mist2_w4w8m9_052110;
  if visitid = 'W04' then output m01;
  else if visitid = 'W08' then output m02;
  else if visitid = 'M09' then output m09;
  else abort;


*** Macro to create a dataset that matches table 1;
%macro baseline_table1(dataset_aua,dataset_qmax,visit);

  proc sort data = &dataset_aua;   by releaseid;
  proc sort data = &dataset_qmax;  by releaseid;
  proc sort data = grou_id;        by releaseid;

  data &visit._combine;
    merge &dataset_aua &dataset_qmax grou_id;
```

```
    by releaseid;

  proc sort data = &visit._combine; by group;

  proc means data = &visit._combine noprint;
    var &visit._aua &visit._qmax;
    by group;
        output out=&visit._meanout;
    title3 "&visit._combine by group";

  data &visit._meanout(drop  =  _TYPE_    _FREQ_    _STAT_  );
    set &visit._meanout;
    if _STAT_ = 'MEAN';

  proc print data = &visit._meanout;
    title3 "&visit._meanout";

  data &visit._meanout_aua  (keep = group &visit._aua)
       &visit._meanout_qmax (keep = group &visit._qmax);
    set &visit._meanout;

  proc transpose data=&visit._meanout_aua  out=&visit._trans_aua   ;
  proc transpose data=&visit._meanout_qmax out=&visit._trans_qmax  ;
/*
  proc print data = &visit._trans_aua;  title3 "&visit._trans_aua";
  proc print data = &visit._trans_qmax; title3 "&visit._trans_qmax";
*/
  data &visit._trans_aua(drop = _LABEL_ rename = (COL1 = aua_100 COL2 = aua_300));
    set &visit._trans_aua;
    if _NAME_ ne 'group';

  data &visit._trans_aua;
    set &visit._trans_aua;
    length visit_char $ 10.;
    visit_char = upcase(substr(_NAME_,1, index(_NAME_,'_')-1));

  data &visit._trans_qmax(drop = _LABEL_ rename = (COL1 = qmax_100 COL2 = qmax_300));
    set &visit._trans_qmax;
    if _NAME_ ne 'group';

  data &visit._trans_qmax;
    set &visit._trans_qmax;
    length visit_char $ 10.;
    visit_char = upcase(substr(_NAME_,1, index(_NAME_,'_')-1));
/*
  proc print data = &visit._trans_aua;  title3 "&visit._trans_aua";
  proc print data = &visit._trans_qmax; title3 "&visit._trans_qmax";
*/
  data &visit._trans_all ;
    merge &visit._trans_aua (drop = _NAME_) &visit._trans_qmax (drop = _NAME_);
    by visit_char;

  proc print data = &visit._trans_all; title3 "&visit._trans_all";

%mend;

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

*** Running the baseline_table1 on the table1 manuscript file;
%baseline_table1(aua_S01,baseline_other,baseline     );
%baseline_table1(aua_M01,M01,M01     );
%baseline_table1(aua_M02,M02,M02     );
%baseline_table1(aua_M03,M03,M03     );
%baseline_table1(aua_M06,M06,M06     );
%baseline_table1(aua_M09,M09,M09     );
%baseline_table1(aua_M12,M12,M12     );

data table1_compare;
  set
```

```
  baseline_trans_all
M01_trans_all
M02_trans_all
M03_trans_all
M06_trans_all
M09_trans_all
M12_trans_all
;
proc print data = table1_compare;
  title3 'table1_combine';

*** Importing the Table 1 Data taken from the primary outcome paper;
data table1_data;
  infile table1 delimiter = ',' MISSOVER DSD  firstobs=1 ls=1080;
  length visit_char $10  ;
  input
  visit_char $ aua_100_stat    aua_300_stat  qmax_100_stat  qmax_300_stat char_stat $
;

proc print data = table1_data;
  title3 'table1_data';

proc sort data = table1_compare; by visit_char;
proc sort data = table1_data; by visit_char;

data combined_table1_dataset;
  merge table1_compare
        table1_data    ;
  by visit_char;
  aua_100    = round (aua_100   ,0.1);
  aua_300    = round (aua_300   ,0.1);
  qmax_100   = round (qmax_100  ,0.1);
  qmax_300   = round (qmax_300  ,0.1);
  diff1 = round((aua_100_stat  - aua_100 ), 0.1);
  diff2 = round((aua_300_stat  - aua_300 ), 0.1);
  diff3 = round((qmax_100_stat - qmax_100), 0.1);
  diff4 = round((qmax_300_stat - qmax_300), 0.1);

    label
aua_100_stat   = "AUASI 100U [Manuscript]"
aua_100        = "AUASI 100U [DSIC]      "
diff1          = "AUASI 100U [Difference]"
aua_300_stat   = "AUASI 300U [Manuscript]"
aua_300        = "AUASI 300U [DSIC]      "
diff2          = "AUASI 300U [Difference]"
qmax_100_stat  = "Qmax 100U  [Manuscript]"
qmax_100       = "Qmax 100U  [DSIC]      "
diff3          = "Qmax 100U  [Difference]"
qmax_300_stat  = "Qmax 300U  [Manuscript]"
qmax_300       = "Qmax 300U  [DSIC]      "
diff4          = "Qmax 300U  [Difference]"
;

*** Outputting the dataset to a csv file to be added to the DSIC;

ods csv file = out_t1;

run;

proc print data = combined_table1_dataset NOOBS label;
       var visit_char aua_100_stat  aua_100  diff1 aua_300_stat  aua_300  diff2  qmax_100_stat
qmax_100  diff3   qmax_300_stat qmax_300  diff4
        ;
       title3 "DSIC Check of Table 1";

run;

ods csv close;
```

```
*******************************************************************************;
*************** Check Table 2 ************************************************;
*******************************************************************************;

proc sort data = aua;
  by releaseid;
  where visitid = 'S01';

data table2_combine;
  merge aua (in = in1) baseline_other (in = in2);
  by releaseid;
  if in2;

proc contents data = table2_combine;

proc sort data = table2_combine;
  by group;

proc freq data = table2_combine noprint;
  tables   raceeth  /out=table2_racefreq ;
  tables    s1marrie /out=table2_marriagefreq ;
  by group;

proc print data = table2_racefreq;
  title3 'table2_racefreq';

proc print data = table2_marriagefreq;
  title3 'table2_marriagefreq';


proc means data = table2_combine noprint;
  var age  bph_baseline total_aua baseline_qmax p_len_baseline  p_vol_baselin msa_baseline;
  by group;
       output out=table2_mean;
  title3 'means of table 2 ';

proc print data = table2_mean;
  title3 'table2_mean';

data table2_com;
  merge table2_racefreq     (keep = group PERCENT raceeth rename = (PERCENT = race_PERCENT) where
= (raceeth=1))
        table2_marriagefreq (keep = group PERCENT s1marrie rename = (PERCENT = marriage_PERCENT)
where = (s1marrie=1))
        table2_mean (drop = _TYPE_     _FREQ_  where = (_STAT_ in ('MEAN')) rename = (age =
age_mean bph_baseline = bph_mean total_aua= aua_mean baseline_qmax=qmax_mean
p_len_baseline=p_len_mean p_vol_baselin=p_vol_mean msa_baseline=msa_mean))
        table2_mean (drop = _TYPE_     _FREQ_  where = (_STAT_ in ('STD')) rename = (age = age_std
bph_baseline = bph_std total_aua= aua_std baseline_qmax=qmax_std p_len_baseline=p_len_std
p_vol_baselin=p_vol_std msa_baseline=msa_std))
        ;
  by group;

proc print data = table2_com;
  title3 'table2_com';

proc transpose data=table2_com (drop = raceeth s1marrie) out=table2_trans   ;

proc print data = table2_trans;
  title3 'table2_trans';

*** Importing the Table 1 Data taken from the primary outcome paper;
data table2_data;
  infile table2 delimiter = ',' MISSOVER DSD  firstobs=1 ls=1080;
  length characteristic $45 visit_char $20  ;
  input
  characteristic $ visit_char $ stat1 stat2 test $
;

proc print data = table2_data;
  title3 'table2_data';
```

```
data table2_data;
  set table2_data;
  sort_order = _n_;
  visit_char = upcase(visit_char);

data table2_trans;  set table2_trans; length visit_char $ 20.; visit_char = upcase(_NAME_);

proc sort data = table2_trans; by visit_char;
proc sort data = table2_data; by visit_char;

data combined_table2_dataset;
  merge table2_trans
        table2_data    ;
  by visit_char;
  COL1    = round (COL1   ,0.1);
  COL2    = round (COL2   ,0.1);
  diff1 = round((stat1  - COL1 ), 0.1);
  diff2 = round((stat2  - COL2 ), 0.1);

    label
stat1   = " 100U [Manuscript]"
COL1    = " 100U [DSIC]      "
diff1   = " 100U [Difference]"
stat2   = " 300U [Manuscript]"
COL2    = " 300U [DSIC]      "
diff2   = " 300U [Difference]"
;

proc sort data = combined_table2_dataset;  by sort_order;

*** Outputting the dataset to a csv file to be added to the DSIC;

ods csv file = out_t2;

run;

proc print data = combined_table2_dataset NOOBS label;
       var characteristic visit_char stat1 col1 diff1 stat2 col2 diff2;
       title3 "DSIC Check of Table 2";

run;

ods csv close;


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


data table3_var(keep =group releaseid age  total_aua baseline_qmax   p_vol_baselin);
  set table2_combine;

/* email from Kathy Hirst Wednesday, April 16, 2014 4:27 PM */
data table3_id(keep = releaseid);
  set AUA_M12;
  if releaseid not in (182 352 417 473 526 596 721 763 774 810  );

proc sort data = table3_var;
  by releaseid;

data table3_1 table3_2;
  merge table3_var (in = in1) table3_id (in = in2);
  by releaseid;
  if in1 and in2 then output table3_1;
  if in1 and not in2 then output table3_2;


proc means data = table3_1 noprint;
  var age  total_aua baseline_qmax   p_vol_baselin;
       output out=table3_1_mean;
```

```
  title3 'means of table3_1 ';

proc means data = table3_2 noprint;
  var age   total_aua baseline_qmax   p_vol_baselin;
       output out=table3_2_mean;
  title3 'means of table3_2 ';

proc print data = table3_1_mean;
  title3 'table3_1_mean';
proc print data = table3_2_mean;
  title3 'table3_2_mean';

proc transpose data=table3_1_mean (where =( _STAT_ in ('MEAN' 'STD')) drop  = _TYPE_    _FREQ_  )
out=table3_1_trans   ;
proc transpose data=table3_2_mean (where =( _STAT_ in ('MEAN' 'STD')) drop  = _TYPE_    _FREQ_  )
out=table3_2_trans   ;

data table3_1_trans(drop = _LABEL_ rename = (col1= mean_108 col2 = std_108 )); set table3_1_trans;
length visit_char $ 10.; visit_char = _name_;
data table3_2_trans(drop = _LABEL_ rename = (col1= mean_26 col2 = std_26   )); set table3_2_trans;
length visit_char $ 10.; visit_char = _name_;

proc print data = table3_2_trans; title3 'table3_2_trans';
proc print data = table3_1_trans; title3 'table3_1_trans';

proc sort data = table3_2_trans (drop = _name_); by visit_char;
proc sort data = table3_1_trans (drop = _name_); by visit_char;

data table3_compare ;
  merge table3_1_trans table3_2_trans;
  by visit_char;

*** Importing the Table 1 Data taken from the primary outcome paper;
data table3_data;
  infile table3 delimiter = ',' MISSOVER DSD  firstobs=1 ls=1080;
  length  visit_char $10  ;
  input
   visit_char $ stat1  stat2 stat3 stat4 test $
;

proc print data = table3_data;
  title3 'table3_data';
data table3_data;
  set table3_data;
  sort_order = _n_;

proc sort data = table3_data; by visit_char;

data combined_table3_dataset;
  merge table3_compare
        table3_data    ;
  by visit_char;
  mean_108  = round (mean_108   ,0.1);
  std_108   = round (std_108   ,0.1);
  mean_26   = round (mean_26    ,0.1);
  std_26    = round (std_26     ,0.1);

  diff1 = round((stat1  - mean_108 ), 0.1);
  diff2 = round((stat2  - std_108  ), 0.1);
  diff3 = round((stat3  - mean_26  ), 0.1);
  diff4 = round((stat4  - std_26   ), 0.1);

  compare_char_108 = strip(put(mean_108,8.1)) || " +- " || strip(put(std_108,8.1)) ;
  compare_char_26  = strip(put(mean_26 ,8.1)) || " +- " || strip(put(std_26 ,8.1)) ;

  stat_char_108 = strip(put(stat1,8.1)) || " +- " || strip(put(stat2,8.1)) ;
  stat_char_26  = strip(put(stat3,8.1)) || " +- " || strip(put(stat4,8.1)) ;

  diff_char_108 = strip(put(diff1,8.1)) || " +- " || strip(put(diff2,8.1)) ;
  diff_char_26  = strip(put(diff3,8.1)) || " +- " || strip(put(diff4,8.1)) ;
```

```
    label
     stat_char_108          = "12-mo completers (108) [Manuscript]"
     compare_char_108       = "12-mo completers (108) [DSIC]      "
     diff_char_108          = "12-mo completers (108) [Difference]"
     stat_char_26          = "Dropouts(26) [Manuscript]"
     compare_char_26       = "Dropouts(26) [DSIC]      "
     diff_char_26          = "Dropouts(26) [Difference]"
;


proc sort data = combined_table3_dataset;  by sort_order;

*** Outputting the dataset to a csv file to be added to the DSIC;

ods csv file = out_t3;

run;

proc print data = combined_table3_dataset NOOBS label;
        var  visit_char stat_char_108 compare_char_108 diff_char_108 stat_char_26 compare_char_26
diff_char_26    ;
        title3 "DSIC Check of Table 3";

run;

ods csv close;

endsas;
```