Dataset Integrity Check for the Complementary and Alternative Medicine for Urological Symptoms (CAMUS) Trial Data Files

Prepared by

RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709-2194
August 18, 2012

Revision History

Version	Author / Title	Date	Comments
1.0	Norma Pugh	August 18, 2012	Original

Table of Contents

1 Standard Disclaimer 1
2 Study Background 1
3 Archived Datasets 2
4 Statistical Methods 2
5 Results 2
6 Conclusion 3
7 References 3
Attachment A: SAS Code 10

Table A: Variables Used to Replicate Table 1, Baseline Characteristics of Participants Included in the Modified Intention-to-Treat Analysis 4
Table B: Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values 5
Table C: Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, cont. 6
Table D: Variables Used to Replicate Table 3, Number of Adverse Events by Treatment Group in the Modified Intention-to-Treat Analysis 7
Table E: Comparison of Values Computed in Integrity Check to Reference Article Table 3 Values 8
Table F: Comparison of Values Computed in Integrity Check to Reference Article Table 3 Values 9

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 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 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

Saw palmetto fruit extracts are widely used for treating lower urinary tract symptoms (LUTS) attributed to benign prostate hyperplasia (BPH). A variety of mechanisms for saw palmetto have been proposed including anti-androgenic, anti-inflammatory, and antiproliferative effects, but none have been conclusively proven. In fact, recent clinical trials have questioned the efficacy of saw palmetto fruit extracts at standard doses (320 mg/d) [1].

The CAMUS Trial is a randomized, double-blind, placebo-controlled trial of increasing doses of saw palmetto fruit extract. The trial, conducted at eleven North American clinical centers, aims to determine the effect of saw palmetto extract (*Serenoa repens*, from saw palmetto berries) at up to 3 times the standard dose on LUTS attributable to BPH. 369 men aged 45 years or older were assigned to receive 1, 2, and then 3 doses (320 mg/d) of saw palmetto extract or placebo, with dose increases at 24 and 48 weeks. The men were followed for a total of 72 weeks [1].

Barry et al report that increasing doses of a saw palmetto fruit extract did not reduce LUTS more than placebo. Additionally, saw palmetto fruit extract was no more effective than placebo for any secondary outcome. No clearly attributable adverse events were identified [1].

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the CAMUS Data folder in the Official Archive. For this replication, variables were taken from both SAS data files, ALL_DERIVED and AESAE, located in the located in the official archive for the study.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Barry et al [1] in The Journal of the American Medical Association in September 2011.

To verify the integrity of the ALL_DERIVED SAS data file housed at the repository, descriptive statistics of baseline characteristics were computed, by treatment group (Tables B, C). To verify the integrity of the AESAE SAS data file housed at the repository, frequencies of adverse events were computed, by treatment group (Tables E, F). The SAS code for our analysis is included in Attachment 1.

5 Results

Table 1 in the publication [1], Baseline Characteristics of Participants Included in the Modified Intention-to-Treat Analysis, reports on baseline characteristics by treatment group. Our Table A lists the variables we used in our replication and Tables B and C 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.

Table 3 in the publication [1], Number of Adverse Events by Treatment Group in the Modified Intention-to-Treat Population, reports frequencies of adverse events by treatment group. Our Table D lists the variables we used in our replication and Tables E and F compare the results calculated from the archived data file to the results published in Table 3. Again, the results of the replication are similar to published results.

6 Conclusion

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

7 References

1. Michael J. Barry, MD, et al for the Complementary and Alternative Medicine for Urological Symptoms (CAMUS) Study Group; **Effect of Increasing Doses of Saw Palmetto Extract on Lower Urinary Tract Symptoms**; The Journal of the American Medical Association; September 2011; Vol 306, No. 12, 1344-1350.

Table A: Variables Used to Replicate Table 1, Baseline Characteristics of Participants Included in the Modified Intention-to-Treat Analysis

Table Variable	Variables Used in Replication
Treatment assignment	drug (1 = saw palmetto extract, 2 = placebo), visitno=3, pool357=1
Age	enroll_age, visitno=3, pool357=1
Race/ethnicity	cm21_arace, cm21_brace5, cm21_brace4, visitno=1, pool357=1
Education	cm21_hedu, visitno=1, pool357=1
AUASI score	auass, visitno=3, pool357=1
BPH Impact Index score	bph, visitno=3, pool357=1
IPSS QOL score	qol, visitno=3, pool357=1
AUA nocturia item	noc, visitno=3, pool357=1
Peak uroflow	cm42_peakflowrt, visitno=3, pool357=1
Postvoid residual	cm42_postvoidres, visitno=3, pool357=1
PSA level	cm41_psaresult, visitno=3, pool357=1
IIEF scale	ef, visitno=3, pool357=1
MSHQ-EjD scale	ejscale, visitno=3, pool357=1
ICSmaleIS score	ics, visitno=3, pool357=1
Jenkins Sleep Dysfunction Scale score	jscale, visitno=3, pool357=1
NIH CPSI Pain scale	pain, visitno=3, pool357=1
NIH CPSI Urinary symptom scale	urin, visitno=3, pool357=1
NIH CPSI QOL scale	qoli, visitno=3, pool357=1

*All variables taken from dataset ALL_DERIVED.

Table B: Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values

Strata = Saw Palmetto Extract and Placebo

Characteristic	Saw Palmetto Extract			Placebo		
	Barry	Integrity Check	Difference	Barry	Integrity Check	Difference
Sample size, No.	176	176	0	181	181	0
Age, y	61.25 (8.72)	61.25 (8.72)	0	60.7 (8.08)	60.7 (8.08)	0
Race/ethnicity, No. (%) Non-Hispanic White Black Hispanic, Latino, or other	145 (82.4) 17 (9.7) 14 (8.0)	145 (82.4) 19 (10.8) 12 (6.8)	0 +2 -2	139 (76.8) 24 (13.3) 18 (9.9)	140 (77.3) 24 (13.3) 17 (9.4)	+1 0 -1
Education, No. (%) <High school High school graduate Some college College graduate Postcollege No response	6 (3.4) 20 (11.4) 26 (14.8) 48 (27.3) 75 (42.6) 1 (0.6)	6 (3.4) 20 (11.4) 26 (14.8) 48 (27.3) 75 (42.6) 1 (0.6)	0	7 (3.9) 18 (9.9) 34 (18.8) 51 (28.2) 67 (37.0) 4 (2.2)	7 (3.9) 18 (9.9) 34 (18.8) 51 (28.2) 67 (37.0) 4 (2.2)	0
AUASI score	14.42 (4.29)	14.42 (4.29)	0	14.69 (4.75)	14.69 (4.75)	0
BPH Impact Index score	3.39 (2.24)	3.39 (2.24)	0	3.71 (2.72)	3.71 (2.72)	0
IPSS QOL score	3.2 (1.2)	3.2 (1.2)	0	3.23 (1.21)	3.23 (1.21)	0
AUA nocturia item	2.09 (1.08)	2.09 (1.08)	0	2.26 (1.13)	2.26 (1.13)	0
Peak uroflow, mL/s	15.03 (7.15)	15.03 (7.15)	0	14.78 (6.71)	14.78 (6.71)	0
Postvoid residual, median (IQR), mL	37.5 (13.5-88.0)	37.5 (13.5-88.0)	0	43.0 (12.0-92.0)	43.0 (12.0-92.0)	0
PSA level, ng/mL	2.20 (1.95)	2.20 (1.95)	0	1.93 (1.59)	1.93 (1.59)	0
IIEF scale	18.79 (10.36)	18.79 (10.36)	0	19.93 (9.43)	19.93 (9.43)	0
MSHQ-EjD scale	10.56 (4.27)	10.56 (4.27)	0	11.18 (4.03)	11.18 (4.03)	0
ICSmaleIS score	3.44 (2.3)	3.44 (2.3)	0	4.17 (3.08)	4.17 (3.08)	0
Jenkins Sleep Dysfunction Scale score	6.95 (4.28)	6.95 (4.28)	0	7.72 (4.93)	7.72 (4.93)	0
NIH CPSI Pain scale, median (IQR) Urinary symptom scale QOL scale	0 (0-2) 4.02 (2.31) 4.45 (2.00)	0 (0-2) 4.02 (2.31) 4.45 (2.00)	0	0 (0-30) 4.27 (2.08) 4.57 (2.24)	0 (0-3) 4.27 (2.08) 4.57 (2.24)	0 (0,-27) 0 (0) 0 (0)

Data are presented as mean (SD) unless otherwise specified.

Table C: Comparison of Values Computed in Integrity Check to Reference Article Table 1 Values, cont.

Strata = Total; P-values

Characteristic	Total			P-value		
	Barry	Integrity Check	Difference	Barry	Integrity Check	Difference
Sample size, No.	357	357	0	n/a		
Age, y	60.97 (8.40)	60.97 (8.40)	0	.54	.54	0
Race/ethnicity, No. (%) Non-Hispanic White Black Hispanic, Latino, or other	284 (79.6) 41 (11.5) 32 (9.0)	285 (79.8) 43 (12.0) 29 (8.1)	+1 +2 -3	.42	.48	+.06
Education, No. (%) <High school High school graduate Some college College graduate Postcollege No response	13 (3.6) 38 (10.6) 60 (16.8) 99 (27.7) 142 (39.8) 5 (1.4)	13 (3.6) 38 (10.6) 60 (16.8) 99 (27.7) 142 (39.8) 5 (1.4)	0	.64	.59	-0.05
AUASI score	14.55 (4.52)	14.55 (4.52)	0	.58	.58	0
BPH Impact Index score	3.55 (2.51)	3.55 (2.50)	0 (-0.01)	.30	.23	
IPSS QOL score	3.21 (1.20)	3.21 (1.20)	0	.83	.83	0
AUA nocturia item	2.17 (1.11)	2.17 (1.11)	0	.14	.14	0
Peak uroflow, mL/s	14.90 (6.92)	14.90 (6.92)	0	.74	.74	0
Postvoid residual, median (IQR), mL	41.0 (13.0-90.0)	41.0 (13.0-92.0)	0 (0,+2)	.88	.86	-0.02
PSA level, ng/mL	2.07 (1.78)	2.07 (1.78)	0	.16	.16	0
IIEF scale	19.38 (9.87)	19.37 (9.90)	-0.01 (+0.03)	.29	.28	-0.01
MSHQ-EjD scale	10.87 (4.16)	10.87 (4.16)	0	.16	.16	0
ICSmaleIS score	3.81 (2.75)	3.81 (2.75)	0	.01	.01	0
Jenkins Sleep Dysfunction Scale score	7.36 (4.62)	7.34 (4.63)	-0.02 (+0.01)	.11	.12	+0.01
NIH CPSI Pain scale, median (IQR) Urinary symptom scale QOL scale	0 (0-2) 4.15 (2.20) 4.51 (2.13)	0 (0-3) 4.15 (2.20) 4.51 (2.13)	0 (0,+1) 0 (0) 0 (0)	.17 .28 .61	.07 .28 .59	-0.1 0 -0.02

Data are presented as mean (SD) unless otherwise specified.
P-values are computed from 2-sample t tests or Wilcoxon rank sum test.

Table D: Variables Used to Replicate Table 3, Number of Adverse Events by Treatment Group in the Modified Intention-to-Treat Analysis

Table Variable	Variables Used in Replication
Type of adverse event	EventType, pool357=1, flag_ae=1

*All variables taken from dataset AESAE.

Table E: Comparison of Values Computed in Integrity Check to Reference Article Table 3 Values

No. of Adverse Events

Type of AE	Stratum = Saw Palmetto Extract			Stratum = Placebo		
	Barry	Integrity Check	Difference	Barry	Integrity Check	Difference
All adverse events	530	530	0	476	477	+1
Musculoskeletal	81	77	-4	72	72	0
Genitourinary	58	58	0	59	60	+1
Upper respiratory tract	54	53	-1	60	61	+1
Gastrointestinal	52	52	0	58	57	-1
Physical injury or trauma	28	28	0	11	10	-1
Oral or dental	26	26	0	14	14	0
Flu-like symptoms	19	19	0	15	15	0
Dermatological	17	17	0	26	28	+2
Increased PSA	15	15	0	15	14	-1
Increased blood pressure	14	14	0	6	6	0
Ophthalmic	11	11	0	11	11	0
Abnormal serum chemistry	11	11	0	10	10	0
Arrhythmia	8	8	0	10	10	0

Table F: Comparison of Values Computed in Integrity Check to Reference Article Table 3 Values

No. of Participants

Type of AE	Stratum = Saw Palmetto Extract			Stratum = Placebo		
	Barry	Integrity Check	Difference	Barry	Integrity Check	Difference
All adverse events	136	136	0	137	137	0
Musculoskeletal	53	52	-1	46	46	0
Genitourinary	41	41	0	42	42	0
Upper respiratory tract	39	39	0	34	35	+1
Gastrointestinal	38	38	0	39	39	0
Physical injury or trauma	24	24	0	10	10	0
Oral or dental	21	21	0	12	12	0
Flu-like symptoms	16	16	0	12	12	0
Dermatological	12	12	0	20	22	+2
Increased PSA	14	14	0	13	13	0
Increased blood pressure	13	13	0	6	6	0
Ophthalmic	8	8	0	9	9	0
Abnormal serum chemistry	11	11	0	7	7	0
Arrhythmia	8	8	0	10	10	0

Attachment A: SAS Code

```
options errorabend;
%include
'\\rcdubuntu01.rtp.rti.org\NIDDK\03_Data_And_Tools\Studies\CAMUS\DCC_Delivery
\20120613\format for CAMUS Main.sas';
/*************************************************************************
**********/
/*
/* Program: R:\05_Users\Norma\CAMUS\table1.sas
/* Author:  Norma Pugh
/* Date:    July 2012
/* Purpose: Replicate table 1 results.
/*
/*************************************************************************
**********/
/* DATA SOURCE */
libname data
'\\rcdubuntu01.rtp.rti.org\NIDDK\03_Data_And_Tools\Studies\CAMUS\DCC_Delivery
\20120615\CAMUS Datasets';

/**********************/
/* ADDITIONAL FORMATS */
/**********************/
proc format;
      value Race 1 = "Non-Hispanic white"
                 2 = "Non-Hispanic black"
                     3 = "Hisp,Latino,Oth";
run;

/****************************/
/* SCREEN/BASELINE DATASETS */
/****************************/
%macro getdata(dset,vnum);
/* Keep visits for the participants randomized and included in the modified
intention-to-treat analysis */
data &dset; set data.all_derived(where=(visitno=&vnum & pool357=1)); run;

/* Restrict to one (last) observation per dataset & Create 'overall'
treatment group */
proc sort data=&dset; by participantid cm01_visdt; run;
data &dset; set &dset; by participantid cm01_visdt; if last.participantid;
run;
data &dset; set &dset; output; drug=99; output; run;
%mend getdata;

%getdata(screen,1);
%getdata(baseline,3);


/***********************************/
/* CREATE DERIVED VARIABLES - SCREEN */
/***********************************/
data screen; set screen;
 /* Race/ethnicity */
```

```
 if      cm21_aRace eq 2 and cm21_bRace5 eq 1 then Race = 1;
 else if cm21_aRace eq 2 and cm21_bRace4 eq 1 then Race = 2;
 else                                              Race = 3;

 /* Education */
 if 1=<cm21_hedu<=5 then edu=cm21_hedu;
 else edu=9;

 format Race Race. Edu Edu.;
run;


/*****************************/
/* REPLICATE ANALYSIS RESULTS */
/*****************************/
proc freq data=baseline; tables drug / list nopct nocum; title'Treatment
Group Counts: Overall'; run;

%macro frq(dset,var);
proc freq data=&dset(where=(&var>.)) noprint; tables drug /
out=denom(keep=drug count rename=(count=denom)); run;
proc freq data=&dset(where=(&var>.)) noprint; tables drug*&var /
out=frqstats(drop=percent); run;
data frqstats; merge frqstats denom; by drug; pct=(count/denom)*100; run;
proc print data=frqstats; title"Frequency Counts: &var"; run;
%mend frq;

proc sort data=screen; by drug; run;
proc sort data=baseline; by drug; run;

%macro mean_(dset,var);
title"Means: &var";
proc means data=&dset n mean std median q1 q3; by drug; var &var; run;

title"T-test p-value: &var";
proc ttest data=&dset(where=(drug in(1,2)));
class drug;
var &var;
run;
%mend mean_;

%macro chisq(dset,var);
title"Chi-square p-value: &var";
proc freq data=&dset(where=(drug in(1,2)));
tables drug*&var / chisq;
run;
%mend chisq;

%macro wilcox(dset,var);
title"Wilcoxon rank sum test p-value: &var";
proc npar1way data=&dset(where=(drug in(1,2))) wilcoxon;
class drug;
var &var;
run;
%mend wilcox;

%mean_(baseline,enroll_age);
```

```
%frq(screen,race);
%chisq(screen,race);
%frq(screen,edu);
%wilcox(screen,edu);
%mean_(baseline,auass);
%mean_(baseline,bph);
%mean_(baseline,qol);
%mean_(baseline,noc);
%mean_(baseline,cm42_peakflowrt);
%mean_(baseline,cm42_postvoidres);
%mean_(baseline,cm41_psaresult);
%mean_(baseline,ef);
%mean_(baseline,ejscale);
%mean_(baseline,ics);
%mean_(baseline,jscale);
%mean_(baseline,pain);
%mean_(baseline,urin);
%mean_(baseline,qoli);
```

```
options errorabend;
%include
'\\rcdubuntu01.rtp.rti.org\NIDDK\03_Data_And_Tools\Studies\CAMUS\DCC_Delivery
\20120613\format for CAMUS Main.sas';
/*************************************************************************
**********/
/*
/* Program: R:\05_Users\Norma\CAMUS\table3.sas
/* Author:  Norma Pugh
/* Date:    July 2012
/* Purpose: Replicate table 3 results.
/*
/*************************************************************************
**********/
/* DATA SOURCE */
libname data
'\\rcdubuntu01.rtp.rti.org\NIDDK\03_Data_And_Tools\Studies\CAMUS\DCC_Delivery
\20120615\CAMUS Datasets';

/*****************************/
/* REPLICATE ANALYSIS RESULTS */
/*****************************/
/* Keep visits for the participants randomized and included in the modified
intention-to-treat analysis */
data ae; set data.aesae(where=(pool357=1)); run;

/* Number of Adverse Events */
data ae1; set ae(where=(flag_ae=1));
if eventtype='Musculoskeletal' then muscu=1; else muscu=0;
if eventtype='Genitourinary' then genit=1; else genit=0;
if eventtype='Upper Respiratory' then ur=1; else ur=0;
if eventtype='Gastrointestinal' then gastro=1; else gastro=0;
if eventtype='Physical Injury/Trauma' then inj=1; else inj=0;
if eventtype='Oral/Dental' then oral=1; else oral=0;
if eventtype='Flu-like symptoms' then flu=1; else flu=0;
if eventtype='Dermatologic' then derm=1; else derm=0;
if eventtype='Elevated PSA' then psa=1; else psa=0;
if eventtype='Elevated Blood Pressure' then bp=1; else bp=0;
if eventtype='Ophthalmic' then ophthal=1; else ophthal=0;
if eventtype='Abnormal serum chemistry' then abnserum=1; else abnserum=0;
if eventtype='Arrhythmia' then arr=1; else arr=0;
run;

%macro frq(var);
proc freq data=ae1; tables drug*&var / list missing; run;
%mend frq;

%frq(EventType);
%frq(Muscu);
%frq(Genit);
%frq(UR);
%frq(Gastro);
%frq(Inj);
%frq(Oral);
%frq(Flu);
%frq(Derm);
%frq(PSA);
```

```
%frq(BP);
%frq(Ophthal);
%frq(Abnserum);
%frq(Arr);

/* Number of participants */
data ae2; set ae(where=(flag_ae=1)); keep participantid drug eventtype; run;

proc sort data=ae2; by participantid eventtype; run;

data ae2; set ae2; by participantid eventtype;
if last.eventtype;
if eventtype='Musculoskeletal' then muscu=1; else muscu=0;
if eventtype='Genitourinary' then genit=1; else genit=0;
if eventtype='Upper Respiratory' then ur=1; else ur=0;
if eventtype='Gastrointestinal' then gastro=1; else gastro=0;
if eventtype='Physical Injury/Trauma' then inj=1; else inj=0;
if eventtype='Oral/Dental' then oral=1; else oral=0;
if eventtype='Flu-like symptoms' then flu=1; else flu=0;
if eventtype='Dermatologic' then derm=1; else derm=0;
if eventtype='Elevated PSA' then psa=1; else psa=0;
if eventtype='Elevated Blood Pressure' then bp=1; else bp=0;
if eventtype='Ophthalmic' then ophthal=1; else ophthal=0;
if eventtype='Abnormal serum chemistry' then abnserum=1; else abnserum=0;
if eventtype='Arrhythmia' then arr=1; else arr=0;
run;

data participants; set ae2; by participantid; if last.participantid; run;

%macro frq(var);
proc freq data=ae2; tables drug*&var / list missing; run;
%mend frq;

proc freq data=participants; tables drug / list missing; run;
%frq(Muscu);
%frq(Genit);
%frq(UR);
%frq(Gastro);
%frq(Inj);
%frq(Oral);
%frq(Flu);
%frq(Derm);
%frq(PSA);
%frq(BP);
%frq(Ophthal);
%frq(Abnserum);
%frq(Arr);
```