Dataset Integrity Check for the HALT PKD Baseline Data Files

Prepared by David Campbell Jr.
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton MD 20705
May 7, 2014

Contents

1 Standard Disclaimer 2

2 Study Background 2

3 Archived Datasets 2

4 Statistical Methods 3

5 Results 3

6 Conclusions 3

7 References 3

Table A: Variables used to replicate Table 1: Table 1 Demographic characteristics of the study population 4

Table B: Comparison of values computed in integrity check to reference article Table 1 values 4

Table C: Variables used to replicate Table 2: Table 2 Baseline characteristics by gender in Study A and Study B. 6

Table D: Comparison of values computed in integrity check to reference article Table 2 values; Study A 7

Table E: Comparison of values computed in integrity check to reference article Table 2 values; Study B 9

Table F: Variables used to replicate Table 3: Table 3 Baseline characteristics in Study A by blood pressure group assignment 11

Table G: Comparison of values computed in integrity check to reference article Table 3 values 12

Attachment A: SAS Code 14

1 Standard Disclaimer

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless NIDDK Repository staff suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

2 Study Background

The HALT PKD (Halt Progression of Polycystic Kidney Disease) studies the efficacy of interruption of the renin-angiotensin- aldosterone system (RAAS) on the progression of cystic disease and on the decline in renal function in autosomal dominant kidney disease (ADPKD); assessed in two multicenter randomized clinical trials targeting different levels of kidney function: (1) early disease defined by GFR >60 mL/min/1.73 m2 (Study A) and moderately advanced disease defined by GFR 25 -60 mL/min/1.73 m2 (Study B). Participants were recruited and enrolled, either to Study A or B, over the first 3½ years. Participants enrolled in Study A were followed for a total of 5 years, or until July 2014. Participants enrolled in Study B were followed until the last clinic visit prior to July 2014 resulting in Study B participants being followed for 5-8 years with the average length of follow up being 6 ½ years. The two concurrent randomized clinical trials differ by eligibility criteria, interventions and outcomes to be studied.

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the HALT PKD data package. For this replication, variables were taken from the “data_baseline_paper” dataset.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Torres et al [1] in Kidney International in December 2011.

To verify the integrity of the dataset, descriptive statistics were computed, by study, sub-study, and gender (Tables B, D, E, and G).

5 Results

Table 1 in the publication [1], Table 1 Demographic characteristics of the study population. 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 within expected results.

Table 2 in the publication [1] Table 2 Baseline characteristics by gender in Study A and Study B. Table C lists the variables that were used in the replication and Tables D and E compare the results calculated from the archived data file to the results published in Table 2. The results of the replication are within expected results.

The data for Table 3 in the publication [1] Table 3 Baseline characteristics in Study A by blood pressure group assignment. Table F lists the variables that were used in the replication and Table G compares the results calculated from the archived data file to the results published in Table 3. The results of the replication are within expected results.

6 Conclusions

The NIDDK repository is confident that the HALT PKD data files to be distributed are within expected results.

7 References

1. Vicente E. Torres, Arlene B. Chapman, Ronald D. Perrone, K. Ty Bae, Kaleab Z. Abebe, James E. Bost, Dana C. Miskulin, Theodore I. Steinman, William E. Braun, Franz T. Winklhofer, Marie C. Hogan, Frederic R. Oskoui, Cass Kelleher, Amirali Masoumi, James Glockner, Neil J. Halin, Diego R. Martin, Erick Remer, Nayana Patel, Ivan Pedrosa, Louis H. Wetzel, Paul A. Thompson, J. Philip Miller, Catherine M. Meyers, Robert W. Schrier. Analysis of baseline parameters in the HALT polycystic kidney disease trials. Kidney International (2012); 81: 577–585.

Table A: Variables used to replicate Table 1: Table 1 Demographic characteristics of the study population

Table Variable	Variables Used in Replication from the "Table 1" Dataset
Study	STUDY
Gender	SEX
Race	RACE
Age at enrollment	AGE
Educational level	EDU
Marital status	MARITAL_STATUS
Employment	EMPLYA EMPLYB EMPLYF EMPLYE OTHER_RETIRE_DISABLED
Diagnosis of ADPKD, age	PKDAGE
Diagnosis due to	DIAGHW
Diagnosis of ADPKD, mode	DIAGNOSIS_METHOD
Diagnosis of hypertension, age	HPBAGE

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

Characteristic	Study A, Standard [Manu-script] (N = 284)	Study A [DSIC] (N = 284)	Study A, Low [Manu-script] (N = 274)	Study A, Low [DSIC] (N = 274)	Study B [Manu-script] (N = 486)	Study B [DSIC] (N = 486)
Gender						
Male (n, %)	143 (50.4)	143 (50.4)	140 (51.2)	140 (51.2)	235 (48.4)	235 (48.4)
Race						
Caucasian (n, %)	258 (90.9)	258 (90.9)	259 (94.5)	259 (94.5)	454 (93.6)	454 (93.6)
African American (n, %)	7(2.5)	7 (2.5)	7 (2.6)	7 (2.6)	12 (2.5)	12 (2.5)
Age at enrollment						
Years (mean±s.d.)	35.9±8.4	35.9±8.4	36.5±8.2	36.5±8.2	48.2±8.3	48.2±8.3
Educational level						
Some high school (n, %)	12(4.2)	12 (4.2)	7 (2.6)	7 (2.6)	2 (0.4)	2 (0.4)
Completed high school (n, %)	33(11.6)	33 (11.6)	31 (11.4)	31 (11.4)	53 (11.0)	53 (11.0)
Some college (n, %)	70(24.7)	70 (24.7)	57 (21.0)	57 (21.0)	117 (24.2)	117 (24.2)
Completed college (n, %)	104 (36.6)	104 (36.6)	111 (40.8)	111 (40.8)	160 (33.1)	160 (33.1)
Graduate studies (n, %)	65(22.9)	65 (22.9)	66 (24.3)	66 (24.3)	152 (31.4)	152 (31.4)
Marital status						
Single (n, %)	82(29.0)	82 (29.0)	80 (29.4)	80 (29.4)	52 (10.7)	52 (10.7)
Married (n, %)	171 (60.4)	171 (60.4)	175 (64.3)	175 (64.3)	363 (74.9)	363 (74.9)

Characteristic	Study A, Standard [Manu-script] (N = 284)	Study A [DSIC] (N = 284)	Study A, Low [Manu-script] (N = 274)	Study A, Low [DSIC] (N = 274)	Study B [Manu-script] (N = 486)	Study B [DSIC] (N = 486)
Divorced/separated (n, %)	2 (9.5)	27 (9.5)	16 (5.9)	16 (5.9)	57 (11.8)	57 (11.8)
Widowed/other (n, %)	3(1.1)	3 (1.1)	1 (0.4)	1 (0.4)	13 (2.6)	13 (2.6)
Employment						
Student (n, %)	25(8.8)	25 (8.8)	27 (9.9)	27 (9.9)	11 (2.3)	11 (2.3)
Homemaker (n, %)	18(6.3)	18 (6.3)	22 (8.0)	22 (8.0)	43 (8.9)	43 (8.9)
Part-time employment	34(12.0)	34 (12.0)	32 (11.7)	32 (11.7)	50 (10.3)	50 (10.3)
Full-time employment	204 (71.8)	204 (71.8)	197 (71.9)	197 (71.9)	342 (70.5)	342 (70.5)
Other/disabled/retired	13(4.6)	13 (4.6)	11 (4.0)	11 (4.0)	60 (12.4)	60 (12.4)
Diagnosis of ADPKD, age						
Years (mean±s.d.)	27.1±9.7	27.1±9.7	28.0±10.3	28.0±10.3	33.1±12.3	33.1±12.2
Diagnosis due to						
Screening (n, %)	113 (39.8)	113 (39.8)	93 (34.2)	93 (34.2)	184 (37.9)	184 (37.9)
Incidental imaging (n, %)	37(13.0)	37 (13.0)	30 (11.0)	30 (11.0)	47 (9.7)	47 (9.7)
Pain (n, %)	42(14.8)	42 (14.8)	34 (12.5)	34 (12.5)	52 (10.7)	52 (10.7)
Hypertension (n, %)	36(12.7)	36 (12.7)	50 (18.4)	50 (18.4)	69 (14.2)	69 (14.2)
Routine physical (n, %)	10(3.5)	10 (3.5)	8 (2.9)	8 (2.9)	26 (5.4)	26 (5.4)
Hematuria (n, %)	15(5.3)	15 (5.3)	25 (9.2)	25 (9.2)	33 (6.8)	33 (6.8)
UTI (n, %)	5(1.8)	5 (1.8)	9 (3.3)	9 (3.3)	9 (1.9)	9 (1.9)
Other (n, %)	26(9.1)	26 (9.1)	23 (8.5)	23 (8.5)	65 (13.4)	65 (13.4)
Diagnosis of ADPKD,						
Ultrasound (n, %)	205 (72.2)	205 (72.2)	195 (71.7)	195 (71.7)	350 (72.2)	350 (72.2)
CT (n, %)	46(16.2)	46 (16.2)	42 (15.4)	42 (15.4)	54 (11.1)	54 (11.1)
MRI (n, %)	17(6.0)	17 (6.0)	16 (5.9)	16 (5.9)	23 (4.7)	23 (4.7)
IVP (n, %)	7(2.5)	7 (2.5)	11 (4.0)	11 (4.0)	31 (6.4)	31 (6.4)
Other (n, %)	9(0.1)	9(3.2)	8 (0.0)	8 (2.94)	27 (0.6)	27 (5.57)
Diagnosis of						
Years (mean±s.d.)	30.2±8.7	30.2±8.7	30.9±9.1	30.9±9.1	36.2±10.6	36.2±10.5

Table C: Variables used to replicate Table 2: Table 2 Baseline characteristics by gender in Study A and Study B.

Table Variable	Variables Used in Replication from the "Table 2" Dataset
Age (years)	AGE
Height (cm)	HGHT_CM
BSA (m2)	BSA
BMI (kg/m2)	BMI
Office systolic BP (mm Hg)	SYS
Office diastolic BP (mmHg)	DIAS
HtTKV (ml/m)	HTTKV
RBF (ml/min per 1.73m2)	RBF_MODIFY
HtTLV (ml/m)	HTLIVER_VOL
Liver cyst volume (ml)	LIVER_CYST_VOLUME
eGFR (ml/min per 1.73m2)	CKD_EPI_EGFR
Serum sodium (mEq/l)	LRSNA
Serum potassium (mEq/l)	LRSK
Urine volume (ml)	UVOL
Urine sodium (mEq/24 h)	USODIUM
Urine potassium (mEq/24 h)	UPOTASSIUM
Urine sodium/potassium ratio	RATIO
Urine aldosterone (mg/24 h)	UALDOS
Urine albumin (mg/24 h)	UALBUM

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

Characteristic	Study A, Male [Manuscript] (N)	Study A, male [DSIC] (N)	Study A, Female [Manuscript] (N)	Study A, Female [DSIC] (N)	Study A, Both [Manuscript]	Study A, Both [DSIC]
Age (years)	35.2±8.1 (283)	35.2±8.1 (283)	37.2±8.4 (275)	37.2±8.4 (275)	36.2±8.3	36.2 ±8.3
Height (cm)	181.0±7.8 (275)	181.0±7.8 (275)	166.3±7.8 (271)	166.3±7.8 (271)	173.7±10.7	173.7±10.7
BSA (m2)	2.1±0.2 (274)	2.1±0.2 (274)	1.8±0.2 (271)	1.8±0.2 (271)	2.0±0.2	2.0±0.2
BMI (kg/m2)	27.6±4.7 (274)	27.6±4.7 (274)	27.2±10.4 (271)	27.2±10.4 (271)	27.4±8.0	27.4±8.0
Office systolic BP (mm Hg)	127.2±14.3 (280)	127.2±14.3 (280)	122.9±14.5 (274)	122.9±14.5 (274)	125.1±14.5	125.1±14.5
Office diastolic BP (mmHg)	80.0±11.4 (280)	80.0±11.4 (280)	78.6±11.8 (274)	78.6±11.8 (274)	79.3±11.6	79.3±11.6
HtTKV (ml/m)	780.7±419.5 (262)	780.7±419.5 (262)	608.9±367.2 (266)	608.9±367.2 (266)	694.1±403.0	694.1 ±403.0
RBF (ml/min per 1.73m2)	665.0±224.3 (130)	665.0±224.3 (130)	610.5±205.4 (138)	610.5±205.4 (138)	636.9±216.1	636.9 ±216.1
HtTLV (ml/m)	1114±402 (265)	1114±402 (265)	1137±513 (269)	1137±513 (269)	1126±461	1126±461
Liver cyst volume (ml)	146.2±703.0 (226)	146.2±703.0 (226)	343.9±795.6 (241)	343.9±795.6 (241)	248.2±757.9	248.2 ±757.9
eGFR (ml/min per 1.73m2)	90.4±17.8 (282)	90.4±17.8 (282)	92.7±17.1 (275)	92.7±17.1 (275)	91.5±17.5	91.5±17.5
Serum sodium (mEq/l)	139.3±2.2 (283)	139.3±2.2 (283)	138.6±7.8 (275)	138.6±7.8 (275)	138.9±5.7	138.9±5.7
Serum potassium (mEq/l)	4.2±0.4 (283)	4.2±0.4 (283)	4.0±0.4 (275)	4.0±0.4 (275)	4.1±0.4	4.1±0.4

Characteristic	Study A, Male [Manuscript] (N)	Study A, male [DSIC] (N)	Study A, Female [Manuscript] (N)	Study A, Female [DSIC] (N)	Study A, Both [Manuscript]	Study A, Both [DSIC]
Urine volume (ml)	2639±1201 (272)	2639±1201 (272)	2457±1150 (265)	2457±1150 (265)	2550±1179	2550±1179
Urine sodium (mEq/24 h)	194.0±75.3 (254)	194.0±75.3 (254)	161.0±78.5 (260)	161.0±78.5 (260)	177.3±78.6	177.3±78.6
Urine potassium (mEq/24 h)	62.9±26.4 (251)	62.9±26.4 (251)	53.5±25.0 (257)	53.5±25.0 (257)	58.1±26.1	58.1±26.1
Urine sodium/ potassium ratio	3.5±1.6 (251)	3.5±1.6 (251)	3.3±1.6 (257)	3.3±1.6 (257)	3.4±1.6	3.4±1.6
Urine aldosterone (mg/24 h)	10.0±5.9 (217)	10.0±5.9 (217)	15.8±12.0 (223)	15.8±12.0 (223)	12.9±9.9	12.9±9.9
Urine albumin (mg/24 h)	40.8±73.2 (254)	40.8±73.2 (254)	42.3±183.6 (260)	42.3±183.6 (260)	41.5±140.2	41.5±140.2

Table E: Comparison of values computed in integrity check to reference article Table 2 values; Study B

Characteristic	Study B, Male [Manuscript] (N)	Study B, male [DSIC] (N)	Study B, Female [Manuscript] (N)	Study B, Female [DSIC] (N)	Study B, Both [Manuscript]	Study B, Both [DSIC]
Age (years)	47.4±8.7 (235)	47.4±8.7 (235)	49.0±7.9 (251)	49.0±7.9 (251)	48.2±8.3	48.2±8.3
Height (cm)	180.3±8.9 (231)	180.3±8.9 (231)	166.4±20.5 (246)	166.4±20.5 (246)	173.1±17.4	173.1±17.4
BSA (m2)	2.1±0.2 (231)	2.1±0.2 (231)	1.8±0.2 (246)	1.8±0.2 (246)	2.0±0.3	2.0±0.3
BMI (kg/m2)	29.0±5.9 (231)	29.0±5.9 (231)	28.2±12.9 (246)	28.2±12.9 (246)	28.6±10.1	28.6±10.1
Office systolic BP (mm Hg)	127.9±15.0 (235)	127.9±15.0 (235)	125.4±15.8 (251)	125.4±15.8 (251)	126.6±15.4	126.6±15.4
Office diastolic BP (mmHg)	80.3±9.7 (234)	80.3±9.7 (234)	76.8±10.9 (251)	76.8±10.9 (251)	78.5±10.5	78.5±10.5
HtTKV (ml/m)	NA		NA		NA	
RBF (ml/min per 1.73m2)	NA		NA		NA	
HtTLV (ml/m)	NA		NA		NA	
Liver cyst volume (ml)	NA		NA		NA	
eGFR (ml/min per 1.73m2)	47.1±11.3 (235)	47.1±11.3 (235)	49.2±12.3 (251)	49.2±12.3 (251)	48.2±11.8	48.2±11.8
Serum sodium (mEq/l)	138.6±11.8 (234)	138.6±11.8 (234)	139.3±2.4 (249)	139.3±2.4 (249)	138.9±8.4	138.9±8.4
Serum potassium (mEq/l)	4.3±0.48 (234)	4.3±0.48 (234)	4.2±0.5 (249)	4.2±0.5 (249)	4.3±0.5	4.3±0.5

Characteristic	Study B, Male [Manuscript] (N)	Study B, male [DSIC] (N)	Study B, Female [Manuscript] (N)	Study B, Female [DSIC] (N)	Study B, Both [Manuscript]	Study B, Both [DSIC]
Urine volume (ml)	2794±1114 (222)	2794±1114 (222)	2541±974 (240)	2541±974 (240)	2662±1050	2662±1050
Urine sodium (mEq/24 h)	202.7±86.6 (211)	202.7±86.6 (211)	153.8±68.1 (224)	153.8±68.1 (224)	177.5±81.3	177.5±81.3
Urine potassium (mEq/24 h)	68.7±28.3 (211)	68.7±28.3 (211)	56.3±23.0 (224)	56.3±23.0 (224)	62.3±26.4	62.3±26.4
Urine sodium/ potassium ratio	3.2±1.2 (211)	3.2±1.2 (211)	3.0±1.5 (224)	3.0±1.5 (224)	3.1±1.3	3.1±1.3
Urine aldosterone (mg/24 h)	10.0±7.7 (173)	10.0±7.7 (173)	10.3±7.6 (190)	10.3±7.6 (190)	10.2±7.6	10.2±7.6
Urine albumin (mg/24 h)	109.1±195.6 (210)	109.1±195.6 (210)	64.9±124.1 (224)	64.9±124.1 (224)	86.3±163.9	86.3±163.9

Table F: Variables used to replicate Table 3: Table 3 Baseline characteristics in Study A by blood pressure group assignment

Table Variable	Variables Used in Replication from the "Table 2" Dataset
Age (years)	AGE
Female	SEX
Height (cm)	HGHT_CM
BSA (m2)	BSA
BMI (kg/m2)	BMI
Office systolic BP (mmHg)	SYS
Office diastolic BP (mmHg)	DIAS
HtTKV (ml/m)	HTTKV
RBF (ml/min per 1.73m2)	RBF_MODIFY
HtTLV (ml/m)	HTLIVER_VOL
Liver cyst volume (ml)	LIVER_CYST_VOLUME
eGFR (ml/min per 1.73m2)	CKD_EPI_EGFR
Serum sodium (mEq/l)	LRSNA
Serum potassium (mEq/l)	LRSK
Urine volume (ml)	UVOL
Urine sodium (mEq/24 h)	USODIUM
Urine potassium (mEq/24 h)	UPOTASSIUM
Urine sodium/potassium ratio	RATIO
Urine aldosterone (mg/24 h)	UALDOS
Urine albumin (mg/24 h)	UALBUM

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

Characteristic	Study A, Standard [Manu-script] (N)	Study A, Stand-ard [DSIC] (N)	Study A, Standard [Manu-script] (Mean±s.d.)	Study A, Standard [DSIC] (Mean ±s.d.)	Study A, Low [Manu-script] (N)	Study A, Low [DSIC] (N)	Study A, Low [Manu-script] (Mean ±s.d.)	Study A, Low [DSIC] (Mean ±s.d.)
Age (years)	284	284	35.9±8.4	35.9 ±8.4	274	274	36.5±8.2	36.5±8.2
Female	284	284	49.6%	49.6%	274	274	48.9%	48.9%
Height (cm)	280	280	173.4±11.5	173.4 ±11.5	266	266	174.0±9.8	174.0±9.8
BSA (m2)	279	279	2.0±0.2	2.0±0.2	266	266	2.0±0.2	2.0±0.2
BMI (kg/m2)	279	279	27.8±10.1	27.8 ±10.1	266	266	27.0±5.1	27.0±5.1
Office systolic BP (mm Hg)	282	282	125.2±14.6	125.2 ±14.6	272	272	125.0±14.5	125.0±14.5
Office diastolic BP (mmHg)	282	282	79.9±11.7	79.9 ±11.7	272	272	78.7±11.5	78.7±11.5
HtTKV (ml/m)	269	269	704.2±406.1	704.2 ±406.1	259	259	683.7 ±400.2	683.7±400.2
RBF (ml/min per 1.73m2)	131	131	623.2±215.0	623.2 ±215.0	137	137	650.0 ±217.2	650.0±217.2
HtTLV (ml/m)	273	273	1128.4 ±380.4	1128.4 ±380.4	261	261	1122.9 ±532.4	1122.9 ±532.4
Liver cyst volume (ml)	241	241	237.1±596.9	237.1 ±596.9	226	226	260.1 ±899.6	260.1±899.6
eGFR (ml/min per 1.73m2)	283	283	91.7±17.8	91.7 ±17.8	274	274	91.4±17.2	91.4±17.2
Serum sodium (mEq/l)	284	284	139.1±2.3	139.1 ±2.3	274	274	138.8±7.8	138.8±7.8

Characteristic	Study A, Standard [Manu-script] (N)	Study A, Stand-ard [DSIC] (N)	Study A, Standard [Manu-script] (Mean±s.d.)	Study A, Standard [DSIC] (Mean ±s.d.)	Study A, Low [Manu-script] (N)	Study A, Low [DSIC] (N)	Study A, Low [Manu-script] (Mean ±s.d.)	Study A, Low [DSIC] (Mean ±s.d.)
Serum potassium (mEq/l)	284	284	4.1±0.4	4.1±0.4	274	274	4.1±0.4	4.1±0.4
Urine volume (ml)	271	271	2577±1223	2577 ±1223	266	266	2522±1133	2522±1133
Urine sodium (mEq/24 h)	260	260	176.4±77.7	176.4 ±77.7	254	254	178.2±79.7	178.2±79.7
Urine potassium (mEq/24 h)	257	257	57.9±24.4	57.9 ±24.4	251	251	58.4±27.8	58.4±27.8
Urine sodium/potassiu m ratio	257	257	3.4±1.6	3.4±1.6	251	251	3.4±1.6	3.4±1.6
Urine aldosterone (mg/24 h)	226	226	13.9±11.1	13.9 ±11.1	214	214	11.9±8.3	11.9±8.3
Urine albumin (mg/24 h)	260	260	34.9±56.6	34.9 ±56.6	254	254	48.3±191.0	48.3±191.0

Attachment A: SAS Code

```
/*************************************************************************************************************************
***Program: /prj/niddk/ims_analysis/HALT_PKD/prog_initial_analysis/HALT_PKD_dsic.sas;
***Programmer: campbeld
***Date Created: 01may14
*************************************************************************************************************************/;

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


options nofmterr;

%include '/prj/niddk/ims_analysis/HALT_PKD/private_orig_data/HALT Transfer 2014_03/Data Directory/Baseline/Final SAS Programs/HALT_format.sas';

libname sasdataa '/prj/niddk/ims_analysis/HALT_PKD/private_orig_data/HALT Transfer 2014_03/Data Directory/Baseline/Final SAS datasets';

data current;
       set sasdataa.current_baseline_halt;
data paper;
       set sasdataa.data_baseline_paper;
       if liver_Cyst_Volume = 0 then liver_Cyst_Volume = .; /*taken from their program*/

proc freq data=paper;
  table study/missing list;

proc sort data=paper;
       by study;

 *****table 1*****;
proc freq data=paper;
  by study;
  tables Sex Race edu marital_status emplya emplyb emplyf emplye diaghw
         diagnosis_method /list;
  table  other_retire_disabled /missing list;
  title2 'table 1';

proc means data=paper;
  by study;
  var age pkdage hpbage;

***table 2***;
proc means data=paper;
  class study_t;
  var age hght_cm bsa bmi sys dias htTKV rbf_modify rbf_total htLiver_vol Liver_Cyst_Volume ckd_epi_egfr
      lrsna lrsk uvol usodium upotassium ratio ualdos ualbum;
  title2 "table 2";

proc means data=paper;
  class study_t sex;
  var age hght_cm bsa bmi sys dias htTKV rbf_modify rbf_total htLiver_vol Liver_Cyst_Volume ckd_epi_egfr
      lrsna lrsk uvol usodium upotassium ratio ualdos ualbum;
  title2 "table 2 (by sex)";
```

```
***table 3***;
proc means data=paper;
  class study;
  where study in (1,2);
  var AGE HGHT_CM BSA BMI SYS DIAS HTTKV RBF_MODIFY HTLIVER_VOL LIVER_CYST_VOLUME CKD_EPI_EGFR
      LRSNA LRSK UVOL USODIUM UPOTASSIUM RATIO UALDOS UALBUM;
  title2 "table 3";

proc freq data=paper;
       by study;
       table sex /missing list;

endsas;
```