Dataset Integrity Check for the Adult-to-Adult Living Donor Liver Transplantation (A2ALL) Core V2 Data Files

Prepared by Nathan Appel and Allyson Mateja
IMS, Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD, 20705
June 21, 2017

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

The Adult-to-Adult Living Donor Liver Transplantation Cohort Study (A2ALL) was a consortium of 9 U.S. liver transplant centers performing adult-to-adult living donor liver transplant (AALDLT) with the primary goal of examining outcomes of AALDLT versus deceased donor liver transplant (DDLT). AALDLT is a relatively new procedure increasingly used at major transplantation centers. Relatively small numbers of cases are performed at any one center and approaches to the patient and donor are too diverse across centers to provide reliable and generalizable information on donor and recipient outcomes from individual centers. Therefore, the consortium was organized to accrue and follow sufficient numbers of patients being considered for and undergoing AALDLT to provide generalizable results from adequately powered studies.

3 Archived Datasets

The SAS data files, as provided by the Data Coordinating Center (DCC), are located in the “Data” folder in the data package. For this replication, variables were taken from the “dnr_basic_all.sas7bdat”, “dnr_eval_rc.sas7bdat”, and “dnr_postdnt_all.sas7bdat” data sets.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Emond, et al. in Liver Transplantation in 2015 [1]. To verify the integrity of the datasets, descriptive statistics were computed.

5 Results

Note that some discrepancies are expected, due to data cleaning of the analysis data set that occurred after the manuscript was published. The images available from the repository do not match those used in the Edmond et al. paper, nor were they used in any publication. However, some data from local reads of images are available in the current data package.

For Table 1 in the publication [1], Characteristics of 388 Living Liver Donors, Table A lists the variables that can be used in the replication. Table B compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are similar, with some discrepancies.

For Table 2 in the publication [1], Donor Laboratory Values and Liver and Spleen Volumes by Time Point: Before Donation and After Donation at M3 and Y1, Table C lists the variables that can be used in the replication. 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, with some discrepancies.

6 Conclusions

The NIDDK repository is confident that the A2ALL data files to be distributed are a true copy of the manuscript data.

7 References

[1] Emond, J. C., Fisher, R. A., Everson, G., Samstein, B., Pomposelli, J. J., Zhao, B., Forney, S., Olthoff, K. M., Baker, T. B., Gillespie, B. W. and Merion, R. M. (2015), Changes in liver and spleen volumes after living liver donation: A report from the adult-to-adult living donor liver transplantation cohort study (A2ALL). Liver Transpl, 21: 151–161. doi:10.1002/lt.24062.

Table A: Variables used to replicate Table 1: Characteristics of 388 Living Liver Donors

Table Variable	**dataset.variable**
Age	dnr_basic_all.dnr_age_ev_rc
Male	dnr_basic_all.dnr_gender_rcx
Female	dnr_basic_all.dnr_gender_rcx
Hispanic	dnr_basic_all.dnr_ethnic_rcx
Non-Hispanic	dnr_basic_all.dnr_ethnic_rcx
White	dnr_basic_all.dnr_race_rcx
African American	dnr_basic_all.dnr_race_rcx
Asian	dnr_basic_all.dnr_race_rcx
Other	dnr_basic_all.dnr_race_rcx
Height	dnr_basic_all.dnr_ht_m_rcx_srs
Weight	dnr_basic_all.dnr_wt_kg_ev_rc
Body mass index	dnr_basic_all.dnr_bmi_ev_rc
Left Lobe Donor	dnr_basic_all.dnr_leftlobe_rcx

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

Characteristic	N Manuscript	N DSIC	Diff.	Mean (SD) or % Manuscript	Mean (SD) or % DSIC	Diff.
Age	388	388	0	37.8 (10.4)	37.8 (10.3)	0 (0.1)
Sex						
Male	189	189	0	48.7	48.7	0
Female	199	199	0	51.3	51.3	0
Ethnicity						
Hispanic	53	55	2	13.7	14.1	0.4
Non-Hispanic	335	332	3	86.3	85.6	0.7
Race						
White	359	354	5	92.5	91.2	1.3
Black	11	10	1	2.8	2.6	0.2
Asian	6	6	0	1.6	1.6	0.0
Other	12	18	6	3.1	4.6	1.5
Height	388	388	0	171.8 (10.2)	171.8 (10.1)	0 (0.1)
Weight	388	388	0	77.9 (14.8)	78.1 (14.9)	0.2 (0.1)
BMI	388	388	0	26.3 (3.9)	26.4 (4.0)	0.1 (0.1)
Left Lobe donor	25	25	0	6.4	6.5	

Characteristic	Range Manuscript	Range DSIC	Diff.
Age	18.2-62.7	18.2-62.7	0-0
Sex			
Male			
Female			
Ethnicity			
Hispanic			
Non-Hispanic			
Race			
White			
Black			
Asian			
Other			
Height	134.6-195.6	134.6-195.6	0-0
Weight	43.1-135.0	43.1-135.0	0-0
BMI	16.4-42.4	16.4-42.4	0-0
Left Lobe donor			

Table C: Variables used to replicate Table 2: Donor Laboratory Values and Liver and Spleen Volumes by Time Point: Before Donation and After Donation at M3 and Y1

Table Variable	**dataset.variable**
Albumin	dnr_eval_rc.dnr_alb_ev_c (Pre-donation) / dnr_postdnt_all.dnr_alb_postdnt_cx (Post Donation)
Bilirubin	dnr_eval_rc.dnr_bili_ev_rc (Pre-donation) / dnr_postdnt_all.dnr_bili_postdnt_cx (Post Donation)
ALT	dnr_eval_rc.dnr_alt_ev_c (Pre-donation) / dnr_postdnt_all.dnr_alt_postdnt_cx (Post Donation)
AST	dnr_eval_rc.dnr_ast_ev_c (Pre-donation) / dnr_postdnt_all.dnr_ast_postdnt_cx (Post Donation)
AP	dnr_eval_rc.dnr_alk_ev_rc (Pre-donation) / dnr_postdnt_all.dnr_alk_postdnt_cx (Post Donation)
INR	dnr_eval_rc.dnr_inr_ev_c (Pre-donation) / dnr_postdnt_all.dnr_inr_postdnt_cx (Post Donation)
Platelet Count	dnr_eval_rc.dnr_plate_ev_c (Pre-donation) / dnr_postdnt_all.dnr_plate_postdnt_cx (Post Donation)
White Blood Count	dnr_eval_rc.dnr_wbc_ev_c (Pre-donation) / dnr_postdnt_all.dnr_wbc_postdnt_cx (Post Donation)
Hemoglobin	dnr_eval_rc.dnr_hgb_ev_c (Pre-donation) / dnr_postdnt_all.dnr_hgb_postdnt_cx (Post Donation)
Liver Volume*	dnr_basic_all.dnr_lvol_predntrcx (Pre-donation) / dnr_postdnt_all.dnr_lvol_postdnt_cx (Post Donation)
Spleen Volume*	N/A
Liver/Spleen Ratio*	N/A

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

Pre-Donation

Characteristic	Mean (SD) Manuscript	Mean (SD) DSIC	Diff.	Range Manuscript	Range DSIC	Diff.
Albumin	4.4 (0.4)	4.4 (0.4)	0	3.5-5.5	3.5-5.5	0-0
Bilirubin	0.7 (0.3)	0.7 (0.3)	0	0.1-2.8	0.1-2.8	0-0
ALT	24.4 (12.5)	24.4 (12.5)	0	4.0-110.0	4.0-110.0	0-0
AST	23.3 (6.6)	23.3 (6.6)	0	11.0-53.0	11.0-53.0	0-0
AP	68.4 (24.8)	67.3 (22.4)	1.1 (2.4)	15.0-197.0	15.0-197.0	0-0
INR	1.00 (0.08)	1.00 (0.08)	0	0.78-1.50	0.78-1.50	0-0
Platelet Count	264.3 (63.0)	264.3 (63.0)	0	126.0-543.0	126.0-543.0	0-0
White Blood Count	6.6 (1.8)	6.6 (1.8)	0	3.1-20.7	3.1-20.7	0-0
Hemoglobin	14.6 (1.4)	14.6 (1.4)	0	9.1-18.4	9.1-18.4	0-0
Liver Volume	1601.0 (327.3)	1568.3 (302.9)	32.7 (24.4)	863.8-3250.8	854.0-2665.0**	9.8-585.8
Spleen Volume*	245.6 (107.3)	--	--	67.1-774.6	--	--
Liver/Spleen Ratio*	7.4 (2.8)	--	--	2.5-22.8	--	--

Characteristic	N Manuscript	N DSIC	Diff.
Albumin	387	387	0
Bilirubin	388	388	0
ALT	388	388	0
AST	387	387	0
AP	388	388	0
INR	382	382	0
Platelet Count	387	387	0
White Blood Count	387	387	0
Hemoglobin	387	387	0
Liver Volume	346	365	19
Spleen Volume*	346	--	--
Liver/Spleen Ratio*	346	--	--

** Note that there was a typo in the original version of the dataset that resulted in a maximum value of 10,616.5 for the pre-donation liver volume. The data set was updated for this value to be corrected to 1023.6 after communications with the DCC.

3 Months Post Donation

Characteristic	Mean (SD) Manuscript	Mean (SD) DSIC	Diff.	Range Manuscript	Range DSIC	Diff.
Albumin	4.1 (0.4)	4.1 (0.4)	0	2.8-5.2	2.8-5.2	0-0
Bilirubin	0.7 (0.4)	0.7 (0.4)	0	0.1-3.6	0.1-3.6	0-0
ALT	29.5 (16.5)	29.6 (16.5)	0	1.5-108.0	1.5-108.0	0-0
AST	29.9 (13.9)	30.0 (13.9)	0.1 (0.0)	13.0-130.0	13.0-130.0	0-0
AP	93.8 (42.9)	93.6 (42.9)	0.2 (0.0)	30.0-385.0	30.0-385.0	0-0
INR	1.05 (0.09)	1.0 (0.09)	0.05 (0.0)	0.89-1.70	0.89-1.7	0-0
Platelet Count	221.6 (67.6)	228.6 (133.7)	7.0 (66.1)	94.0-660.0	94.0-2113.0	0-1453.0
White Blood Count	6.6 (1.7)	6.5 (1.7)	0.1 (0.0)	0.9-14.0	0.9-14.0	0-0
Hemoglobin	13.7 (1.8)	13.7 (1.8)	0	5.4-17.3	5.4-17.3	0-0
Liver Volume	1241.2 (257.1)	1240.5 (271.6)	0.7 (14.5)	790.0-2024.1	760.0-2034.0	30.0-9.9
Spleen Volume*	314.3 (136.4)	--	--	77.6-842.2	--	--
Liver/Spleen Ratio*	4.6 (1.8)	--	--	1.7-11.2	--	--

Characteristic	N Manuscript	N DSIC	Diff.
Albumin	265	264	1
Bilirubin	272	271	1
ALT	272	271	1
AST	272	271	1
AP	271	270	1
INR	254	253	1
Platelet Count	270	269	1
White Blood Count	269	268	1
Hemoglobin	268	267	1
Liver Volume	182	226	44
Spleen Volume*	182	--	--
Liver/Spleen Ratio*	182	--	--

1 Year Post Donation

Characteristic	Mean (SD) Manuscript	Mean (SD) DSIC	Diff.	Range Manuscript	Range DSIC	Diff.
Albumin	4.2 (0.4)	4.2 (0.4)	0	2.9-5.2	2.9-5.2	0-0
Bilirubin	0.8 (0.3)	0.8 (0.3)	0	0.2-2.6	0.2-2.6	0-0
ALT	25.2 (13.5)	25.2 (13.3)	0 (0.2)	6.0-92.0	6.0-92.0	0-0
AST	26.1 (11.0)	26.1 (10.9)	0 (0.1)	11.0-109.0	11.0-109.0	0-0
AP	74.1 (26.3)	74.3 (26.2)	0.2 (0.1)	16.0-186.0	16.0-186.0	0-0
INR	1.02 (0.09)	1.0 2 (0.09)	0	0.70-1.50	0.7-1.5	0-0
Platelet Count	214.4 (63.7)	213.9 (63.2)	0.5 (0.5)	3.6-708.0	3.6-708.0	0-0
White Blood Count	6.6 (1.7)	6.6 (1.7)	0	3.4-16.4	3.4-16.4	0-0
Hemoglobin	14.4 (1.5)	14.4 (1.5)	0	9.4-18.0	9.4-18.0	0-0
Liver Volume*	1440.5 (274.0)	--	--	936.3-2122.0	--	--
Spleen Volume*	323.6 (181.4)	--	--	82.5-1171.7	--	--
Liver/Spleen Ratio*	5.5 (2.6)	--	--	1.2-15.9	--	--

Characteristic	N Manuscript	N DSIC	Diff.
Albumin	192	196	4
Bilirubin	197	201	4
ALT	195	199	4
AST	197	201	4
AP	196	200	4
INR	186	190	4
Platelet Count	92	196	104
White Blood Count	194	198	4
Hemoglobin	192	196	4
Liver Volume*	90	--	--
Spleen Volume*	90	--	--
Liver/Spleen Ratio*	90	--	--

*Note that the variables for Liver Volume at 1 Year Post Donation and Spleen Volume at all three time points are not available in the data package. As a result, the Liver/Spleen Ratio cannot be calculated.

Appendix A: SAS Code

```
*** A2ALL Emond paper;
*** Programmer: Nathan Appel;
*** Date: December 5, 2016;

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

proc format;
       value genderf 1 = 'M'
                     2 = 'F';

       value meldf 6-10  = '6-10'
                   11-20 = '11-20'
                   21-30 = '21-30'
                   31-40 = '31-40';

options nofmterr mprint source2;

***FIN;
PROC IMPORT
  DATAFILE='/prj/niddk/ims_analysis/A2ALL/private_created_data/A2ALL_spleen_paper_study_IDs.xls'
  DBMS=xls
  OUT=a2all_spleen;
  STARTROW=2;
  STARTCOL=A;
  NAMEROW=1; ENDNAMEROW=1;
RUN;


libname sas_data "/prj/niddk/ims_analysis/A2ALL/private_orig_data/A2ALL Core Datav2";
libname out "/prj/niddk/ims_analysis/A2ALL/private_created_data/";

data dnr_basic_all;
       set sas_data.dnr_basic_all;
run;

data rcp_basic_all;
       set sas_data.rcp_basic_all;
run;

proc contents data = dnr_basic_all; run;

proc contents data = rcp_basic_all; run;
```

```
proc contents data = a2all_spleen; run;

proc freq data = a2all_spleen;
       table A2ALL_ID Site VAR2;
run;

data donors_spleen;
       set a2all_spleen;
       if substr(A2ALL_ID,1,1) = 'D' then output donors_spleen;
run;

proc freq data = donors_spleen;
       table A2ALL_ID;
run;

data donors_spleen;
       set donors_spleen (rename = (A2ALL_ID = dnr_ID));
run;

proc contents data = donors_spleen;   run;

proc sort data = donors_spleen;
       by dnr_ID;
run;

proc sort data = dnr_basic_all;
       by dnr_ID;
run;

data donors_spleen;
       merge donors_spleen (in = in1 keep = dnr_ID)
             dnr_basic_all (in = in2);
       by dnr_ID;
       if (dnr_ht_m_rcx_srs ne .) then dnr_ht_m_rcx_srs_cm = dnr_ht_m_rcx_srs*100;
       if (dnr_race_rcx in (1 4 6 9)) then out_race = 9;
       else out_race = dnr_race_rcx;
       if in1 and in2 then output donors_spleen;
run;

proc contents data = donors_spleen; run;

proc means data = donors_spleen;
       var dnr_age_ev_rc dnr_ht_m_rcx_srs_cm dnr_wt_kg_ev_rc dnr_bmi_ev_rc;

proc freq data = donors_spleen;
       table dnr_gender_rcx dnr_ethnic_rcx out_race dnr_leftlobe_rcx;

data dnr_eval_rc;
       set sas_data.dnr_eval_rc;
```

```
run;

proc sort data = dnr_eval_rc;
       by dnr_ID;
run;

data dnr_eval_rc_spleen;
       merge donors_spleen (in = in1 keep = dnr_ID Dnr_lvol_predntrcx)
             dnr_eval_rc   (in = in2);
       by dnr_ID;

       if in1 and in2 then output dnr_eval_rc_spleen;
run;

proc contents data = dnr_eval_rc_spleen; run;

proc freq data=dnr_eval_rc_spleen;
       tables Dnr_lvol_predntrcx;

proc print data = dnr_eval_rc_spleen;
       var dnr_ID;
       where Dnr_lvol_predntrcx=10616.5 ;

data dnr_eval_rc_spleen;
       set dnr_eval_rc_spleen;
       if dnr_ID = 'D1901' then Dnr_lvol_predntrcx=1023.6;

data out.dnr_eval_rc_spleen;
       set dnr_eval_rc_spleen;

proc means data = dnr_eval_rc_spleen;
       var dnr_alb_ev_c dnr_bili_ev_rc dnr_alt_ev_c dnr_ast_ev_c dnr_alk_ev_rc dnr_inr_ev_c dnr_plate_ev_c dnr_wbc_ev_c dnr_hgb_ev_c
Dnr_lvol_predntrcx;
run;

data dnr_postdnt_all;
       set sas_data.dnr_postdnt_all;
run;

proc sort data = dnr_postdnt_all;
       by dnr_ID;
run;

data dnr_postdnt_all_spleen;
       merge donors_spleen (in = in1 keep = dnr_ID)
             dnr_postdnt_all   (in = in2);
       by dnr_ID;

       if in1 and in2 then output dnr_postdnt_all_spleen;
```

```
run;

proc freq data = dnr_postdnt_all_spleen;
       tables dnr_timept_postdnt;


proc contents data = dnr_postdnt_all_spleen; run;

proc means data = dnr_postdnt_all_spleen;
       var dnr_alb_postdnt_cx dnr_bili_postdnt_cx dnr_alt_postdnt_cx  dnr_ast_postdnt_cx dnr_alk_postdnt_cx dnr_inr_postdnt_cx
dnr_plate_postdnt_cx dnr_wbc_postdnt_cx dnr_hgb_postdnt_cx dnr_lvol_postdnt_cx;
       where dnr_timept_postdnt = "month3";
       title3 "month3";
run;

proc means data = dnr_postdnt_all_spleen;
       var dnr_alb_postdnt_cx dnr_bili_postdnt_cx dnr_alt_postdnt_cx  dnr_ast_postdnt_cx dnr_alk_postdnt_cx dnr_inr_postdnt_cx
dnr_plate_postdnt_cx dnr_wbc_postdnt_cx dnr_hgb_postdnt_cx;
       where dnr_timept_postdnt = "year1";
       title3 "year1";
run;
```