

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

Prepared by Allyson Mateja

IMS, Inc.

3901 Calverton Blvd, Suite 200 Calverton, MD, 20705

October 3, 2016

## 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: Recipient Characteristics at Transplant.....	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	5
Attachment A: SAS Code.....	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 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 (DDLTL). 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 “outcomes” data set.

## 4 Statistical Methods

Analyses were performed to duplicate results for the data published by Olthoff, et al. in *Annals of Surgery* 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.

For Table 1 in the publication [1], Recipient Characteristics at Transplant, 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 close, 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] Olthoff, K.M., Smith, A.R., Abecassis, M., Baker, T., Emond, J.C., Berg, C.L., Beil, C.A., Burton, J.R., Fisher, R.A., Freise, C.E., Gillespie, B.W., Grant, D.R., Humar, A., Kam, I., Merion, R.M., Pomfret, E.A., Samstein, B., Shaked, A. Defining Long-term Outcomes With Living Donor Liver Transplantation in North America. *Annals of Surgery* (2015); 262 (3): 465-475.

**Table A:** Variables used to replicate Table 1: Recipient Characteristics at Transplant

<b>Table Variable</b>	<b>Variable in “outcomes” Dataset</b>
DDLT vs. LLDT	LDLT_rcx
Age	rcp_age_tx_rcx
Female	rcp_gender_rcx
Hispanic	hispanic
White	white
Black	black
Asian	asian
Other	other_race
Body mass index	rcp_bmi_tx_rcx
Acute Liver Failure	dgn_acutefail
Alcohol-related cirrhosis	dgn_alccirr
Autoimmune hepatitis	dgn_autoimmhеп
Cryptogenic cirrhosis	dgn_cryptocirr
HBV	dgn_hbv
HCC	dgn_hcc
HCV	dgn_hcv
Hemochromatosis	dgn_hemochrom
Other metabolic liver disease	dgn_metabolic
Malignancy other than HCC	dgn_nonhccmalig
PBC	dgn_pbc
PSC	dgn_psc
Other diagnosis	dgn_oth
MELD at evaluation	rcp_meld_ev_rc
MELD at transplant	rcp_meld_tx_rcx
ICU	icu
Hospitalized not in ICU	hosp
Not hospitalized	nothosp
Ventilator	rcp_vent_tx_rcx
Ascites	rcp_ascites_tx_rc
Dialysis	rcp_dialy_tx_rcx
Duration of recipient surgery (hr)	rcp_skintime_io_rc
Total ischemia time (min)*	N/A
PRBCs	rcp_prbc_io_rc
Recipient ICU LOS	rcp_hosp_icuday_rcx
Recipient total LOS	rcp_hosp_dc_rcx, rcp_txp_dt_rcx

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

	DDLT Manuscript (n=464)	DDLT DSIC (n=464)	Diff. (n=0)						
Characteristic	N Manuscript	N DSIC	Diff.	Mean (SD) or Frequency Manuscript	Mean (SD) or Frequency DSIC	Diff.	Range or % Manuscript	Range or % DSIC	Diff.
Age	463	463	0	52.08 (10.49)	52.08 (10.49)	0 (0)	18-74	18-74	0-0
Female	464	464	0	182	182	0	39%	39%	0
Hispanic	463	463	0	87	88	1	19%	19%	0
Race	464	464	0						
White				390	384	6	84%	83%	1
Black				33	32	1	7%	7%	0
Asian				17	18	1	4%	4%	0
Other				24	30	6	5%	6%	1
Body mass index	412	420	8	26.78 (5.01)	26.73 (4.97)	0.05 (0.04)	13-50	13-50	0-0
Diagnosis	464	464	0						
Acute Liver Failure				19	19	0	4%	4%	0
Alcohol-related cirrhosis				86	86	0	19%	19%	0
Autoimmune hepatitis				20	20	0	4%	4%	0
Cryptogenic cirrhosis				53	53	0	11%	11%	0
HBV				12	12	0	3%	3%	0
HCC				98	98	0	21%	21%	0
HCV				210	210	0	45%	45%	0
Hemochromatosis				3	3	0	1%	1%	0
Other metabolic liver disease				16	16	0	3%	3%	0
Malignancy other than HCC				7	7	0	2%	2%	0
PBC				12	12	0	3%	3%	0
PSC				61	61	0	13%	13%	0
Other diagnosis				21	21	0	5%	5%	0
MELD at evaluation	452	452	0	16.77 (6.61)	16.82 (6.60)	0.05 (0.01)	6-40	6-40	0-0
6-10				67	66	1	15%	15%	0

	DDLT Manuscript (n=464)	DDLT DSIC (n=464)	Diff. (n=0)						
Characteristic	N Manuscript	N DSIC	Diff.	Mean (SD) or Frequency Manuscript	Mean (SD) or Frequency DSIC	Diff.	Range or % Manuscript	Range or % DSIC	Diff.
11-10				267	266	1	59%	59%	0
21-30				97	99	2	21%	22%	1
31-40				21	21	0	5%	5%	0
MELD at transplant	440	440	0	20.42 (8.92)	20.43 (8.92)	0.01 (0)	6-40	6-40	0-0
6-10				52	52	0	12%	12%	0
11-20				201	201	0	46%	46%	0
21-30				118	118	0	27%	27%	0
31-40				69	69	0	16%	16%	0
Medical condition at transplant	462	462	0						
ICU				51	51	0	11%	11%	1
Hospitalized not in ICU				70	72	2	15%	16%	1
Not hospitalized				341	339	2	74%	73%	1
Comorbidities									
Ventilator	461	461	0	28	28	0	6%	6%	0
Ascites	455	455	0	284	285	1	62%	63%	1
Dialysis	457	457	0	25	25	0	5%	5%	0
Perioperative Characteristics	N Manuscript	N DSIC	Diff.	Median Manuscript	Median DSIC	Diff.	IQR Manuscript	IQR DSIC	Diff.
Duration of recipient surgery (hr)	421	421	0	5.78	5.78	0	5-7	5-7	0-0
Total ischemia time (min)*	442	-	-	486.5	-	-	364-600	-	-
PRBCs	439	439	0	6	6	0	3-11	3-11	0-0
Recipient ICU LOS	370	370	0	2	2	0	1-5	1-5	0-0
Recipient total LOS	407	407	0	10	10	0	7-17	7-17	0-0

	LDLT Manuscript (n=963)	LDLT DSIC (n=963)	Diff. (n=0)						
Characteristic	N Manuscript	N DSIC	Diff.	Mean (SD) or Frequency Manuscript	Mean (SD) or Frequency DSIC	Diff.	Range or % Manuscript	Range or % DSIC	Diff.
Age	963	963	0	51.37 (11.48)	51.37 (11.48)	0 (0)	18-76	18-76	0-0
Female	963	963	0	408	408	0	42%	42%	0
Hispanic	963	963	0	126	123	3	13%	13%	0
Race	963	963	0						
White				877	864	11	91%	90%	1
Black				29	30	1	3%	3%	0
Asian				31	31	0	3%	3%	0
Other				26	38	12	3%	4%	1
Body mass index	919	921	2	26.54 (5.26)	26.53 (5.25)	0.01 (0.01)	15-55	15-55	0-0
Diagnosis	963	963	0						
Acute Liver Failure				24	24	0	2%	2%	0
Alcohol-related cirrhosis				155	155	0	16%	16%	0
Autoimmune hepatitis				63	63	0	7%	7%	0
Cryptogenic cirrhosis				80	80	0	8%	8%	0
HBV				28	28	0	3%	3%	0
HCC				154	154	0	16%	16%	0
HCV				339	339	0	35%	35%	0
Hemochromatosis				10	10	0	1%	1%	0
Other metabolic liver disease				21	21	0	2%	2%	0
Malignancy other than HCC				26	26	0	3%	3%	0
PBC				81	81	0	8%	8%	0
PSC				162	163	1	17%	17%	0
Other diagnosis				90	90	0	9%	9%	0
MELD at evaluation	538	538	0	14.55 (6.0)	14.51 (5.97)	0.04 (0.03)	6-40	6-40	0-0
6-10				130	130	0	24%	24%	0
11-10				348	350	2	65%	65%	0
21-30				44	43	1	8%	8%	0



	LDLT Manuscript (n=963)	LDLT DSIC (n=963)	Diff. (n=0)						
Characteristic	N Manuscript	N DSIC	Diff.	Mean (SD) or Frequency Manuscript	Mean (SD) or Frequency DSIC	Diff.	Range or % Manuscript	Range or % DSIC	Diff.
31-40				16	15	1	3%	3%	0
MELD at transplant	935	935	0	15.47 (5.9)	15.47 (5.89)	0 (0.01)	6-40	6-40	0-0
6-10				169	169	0	18%	18%	0
11-20				614	614	0	66%	66%	0
21-30				132	132	0	14%	14%	0
31-40				20	20	0	2%	2%	0
Medical condition at transplant	567	567	0						
ICU				9	9	0	2%	2%	0
Hospitalized not in ICU				35	38	3	6%	7%	1
Not hospitalized				523	520	3	92%	92%	0
Comorbidities									
Ventilator	959	959	0	12	12	0	1%	1%	0
Ascites	567	567	0	260	260	0	46%	46%	0
Dialysis	957	957	0	7	7	0	1%	1%	0
Perioperative Characteristics	N Manuscript	N DSIC	Diff.	Median Manuscript	Median DSIC	Diff.	IQR Manuscript	IQR DSIC	Diff.
Duration of recipient surgery (hr)	533	533	0	7.57	7.57	0	7-9	7-9	0-0
Total ischemia time (min)*	847	-	-	98	-	-	71-140	-	-
PRBCs	557	557	0	4	4	0	2-8	2-8	0-0
Recipient ICU LOS	915	915	0	2	2	0	1-3	1-3	0
Recipient total LOS	945	945	0	10	10	0	7-15	7-15	0-0

\*Note that the variable used to calculate Total Ischemia Time in the manuscript is not included in the current version of the data package.

## Attachment A: SAS Code

```
*** A2ALL Olthoff paper;
*** Programmer: Allyson Mateja;
*** Date: September 7, 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;

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

data outcomes;
    set sas_data.outcomes;

data rcp_basic_rc;
    set sas_data.rcp_basic_rc;

data rcp_dnreval_enroll_all;
    set sas_data.rcp_dnreval_enroll_all;

data rcp_eval_rc;
    set sas_data.rcp_eval_rc;

proc contents data = outcomes;

proc sort data = outcomes;
    by ID;

data outcomes;
    set outcomes;
    if ID not in ('R2892', 'R2905', 'R3498');

proc freq data = outcomes;
    tables DDLT_rc DDLT_rc_srs LDLT_rc LDLT_rc_srs LDLT_rcx;

proc sort data = outcomes;
    by ID;

proc sort data = rcp_basic_rc;
    by ID;
```

```

proc sort data = rcp_dnreval_enroll_all;
  by ID;

proc sort data = rcp_eval_rc;
  by ID;

data outcomes;
  length group $4.;
  set outcomes;
  if LDLT_rcx = 0 then group = 'DDLT';
  else if LDLT_rcx = 1 then group = 'LDLT';
  if rcp_dialy_tx_rcx = 998 then rcp_dialy_tx_rcx = .;
  if rcp_ascites_tx_rc = 998 then rcp_ascites_tx_rc = .;
  length_hours = rcp_skintime_io_rc/60;
  total_length_of_stay = rcp_hosp_dc_rcx - rcp_txp_dt_rcx;
  age55 = rcp_age_tx_rcx - 55;
  if age55 < 0 then age55 = 0;

proc sort data = outcomes;
  by group;

proc means data = outcomes;
  var rcp_age_tx_rcx;
  class group;
  title3 'Table 1 - Age';

proc freq data = outcomes;
  tables rcp_gender_rcx;
  format rcp_gender_rcx genderf.;
  by group;
  title3 'Table 1 - Female';

proc freq data = outcomes;
  tables hispanic ;
  by group;
  title3 'Table 1 - Hispanic';

proc freq data = outcomes;
  tables white black asian other_race;
  by group;
  title3 'Table 1 - Race';

proc means data = outcomes;
  var rcp_bmi_tx_rcx;
  class group;
  title3 'Table 1 - BMI';

proc freq data = outcomes;
  tables dgn_acutefail;
  by group;
  title3 'Table 1 - Diagnosis, Acute Liver Failure';

proc freq data = outcomes;
  tables dgn_alccirr;
  by group;
  title3 'Table 1 - Diagnosis, Alcohol-related cirrhosis';

```

```

proc freq data = outcomes;
    tables dgn_autoimmhep;
    by group;
    title3 'Table 1 - Diagnosis, Autoimmune Hepatitis';

proc freq data = outcomes;
    tables dgn_cryptocirr;
    by group;
    title3 'Table 1 - Diagnosis, Cryptogenic cirrhosis';

proc freq data = outcomes;
    tables dgn_hbv;
    by group;
    title3 'Table 1 - Diagnosis, HBV';

proc freq data = outcomes;
    tables dgn_hcc;
    by group;
    title3 'Table 1 - Diagnosis, HCC';

proc freq data = outcomes;
    tables dgn_hcv;
    by group;
    title3 'Table 1 - Diagnosis, HCV';

proc freq data = outcomes;
    tables dgn_hemochrom;
    by group;
    title3 'Table 1 - Diagnosis, Hemochromatosis';

proc freq data = outcomes;
    tables dgn_metabolic;
    by group;
    title3 'Table 1 - Diagnosis, Other metabolic liver disease';

proc freq data = outcomes;
    tables dgn_nonhccmalig;
    by group;
    title3 'Table 1 - Diagnosis, Malignancy other than HCC';

proc freq data = outcomes;
    tables dgn_pbc;
    by group;
    title3 'Table 1 - Diagnosis, PBC';

proc freq data = outcomes;
    tables dgn_psc;
    by group;
    title3 'Table 1 - Diagnosis, PSC';

proc freq data = outcomes;
    tables dgn_oth;
    by group;
    title3 'Table 1 - Diagnosis, Other Diagnosis';

```

```

proc means data = outcomes;
  var rcp_meld_ev_rc;
  class group;
  title3 'Table 1 - MELD at evaluation';

proc freq data = outcomes;
  tables rcp_meld_ev_rc ;
  format rcp_meld_ev_rc meldf.;
  by group;

proc means data = outcomes;
  var rcp_meld_tx_rcx;
  class group;
  title3 'Table 1 - MELD at transplant';

proc freq data = outcomes;
  tables rcp_meld_tx_rcx ;
  format rcp_meld_tx_rcx meldf.;
  by group;

proc freq data = outcomes;
  tables icu*hosp*nothosp /list ;
  by group;
  title3 'Table 1 - Medical condition at transplant';

proc freq data = outcomes;
  tables rcp_vent_tx_rcx;
  by group;
  title3 'Table 1 - Comorbidities, Ventilator';

proc freq data = outcomes;
  tables rcp_ascites_tx_rc;
  by group;
  title3 'Table 1 - Comorbidities, Ascites';

proc freq data = outcomes;
  tables rcp_dialy_tx_rcx;
  by group;
  title3 'Table 1 - Comorbidities, Dialysis';

proc means data = outcomes n median p25 p75;
  var length_hours;
  class group;
  title3 'Table 1 - Duration of recipient surgery';

proc means data = outcomes n median p25 p75;
  var rcp_prbc_io_rc;
  class group;
  title3 'Table 1 - PRBCs';

proc means data = outcomes n median p25 p75;
  var rcp_hosp_icuday_rcx;
  class group;
  title3 'Table 1 - Recipient ICU LOS (d)';

proc means data = outcomes n median p25 p75;

```

```
var total_length_of_stay;  
class group;  
title3 'Table 1 - Recipient total LOS (d)';
```