Dataset Integrity Check for Acute Liver Failure Study Group (ALFSG) Final Public Use Data Files

Prepared by NIDDK-CR
May 1, 2023

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 Acute Liver Failure Study Group (ALFSG) was a clinical research network that collected data and biospecimens from participants with acute liver failure (ALF) and acute liver injury (ALI). ALF is defined as a severe form of acute liver injury characterized by rapid onset over days and weeks in the absence of underlying cirrhosis, leading to abnormal coagulation with prolongation of the prothrombin time (INR ≥ 1.5) in the presence of altered mentation. ALI is defined as a severe acute liver injury with an INR ≥ 2.0 and no encephalopathy (altered mental functioning, drowsiness, or coma). The primary objectives for the Multi-Center Trial to Study Adult Acute Liver Failure (AALF) were to collect clinical and epidemiological data as well as biospecimens (serum, plasma, urine, tissue, DNA) from participants with ALF, ALI, or those with coagulopathy but did not reach the threshold of encephalopathy.

3 Archived Datasets

All data files, as provided by the Data Coordinating Center (DCC), are located in the ALFSG folder in the data package. For this replication, variables were taken from the “form00.sas7bdat”, “form01.sas7bdat”, “form02.sas7bdat”, “form08.sas7bdat”, and “form45.sas7bdat” datasets.

4 Statistical Methods

Analyses were performed to replicate results for the data in the publication by Stravitz et al. [1]. To verify the integrity of the data, only descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Demographics and Baseline Clinical Features of the Study Population Based on Etiology, Table A lists the variables that were used in the replication, and Tables B1-B2

compare the results calculated from the archived data files to the results in Table 1. The results of the replication are within expected variation of the published results.

6 Conclusions

The NIDDK Central Repository is confident that the ALFSG data files to be distributed are a true copy of the study data.

7 References

[1] Stravitz RT, Fontana RJ, Meinzer C, Durkalski-Mauldin V, Hanje AJ, Olson J, Koch D, Hamid B, Schilsky ML, McGuire B, Ganger D, Liou I, Karvellas CJ, Rule JA, Lisman T, Clasen K, Reuben A, Cripps M, Lee WM. Coagulopathy, Bleeding Events, and Outcome According to Rotational Thromboelastometry in Patients With Acute Liver Injury/Failure. Hepatology, 74(2), 937-949, August 2021. doi: https://doi.org/10.1002/hep.31767

Table A: Variables used to replicate Table 1 – Demographics and Baseline Clinical Features of the Study Population Based On Etiology

Table Variable	**dataset.variable**
Age	form45.zvisitnm form01.age form02.age
Gender	form00.sex form45.zvisitnm
Race	form00.white form45.zvisitnm
International normalized ratio (INR)	form08.inr form45.zvisitnm
Platelets	form08.platelet_cnt form45.zvisitnm
Bilirubin	form08.bilirubin form45.zvisitnm
Creatinine	form08.creat form45.zvisitnm

Table B1: Comparison of values computed in integrity check to reference article Table 1 (ALI)

Characteristic	Acute Liver Injury (ALI)					
	Publication: Non-APAP (n=41)	**DSIC: Non-APAP (n=42)**	**Diff. (n=1)**	**Publication: APAP (n=49)**	**DSIC: APAP (n=48)**	**Diff. (n=1)**
Age (years)	48.0 ± 26.0	47.3 ± 16.1	0.7 ± 9.9	33.0 ± 19.0	35.2 ± 12.5	2.2 ± 6.5
Gender (% female)	18 (43.9)	19 (45.2)	1 (1.3)	34 (69.4)	33 (68.7)	1 (0.7)
Race (% White)	27 (65.9)	27 (64.3)	0 (1.6)	36 (73.5)	36 (75.0)	0 (1.5)
INR	2.4 ± 1.0	-	-	3.1 ± 1.9	0.21 ± 1.3	2.89 ± 0.6
Platelets ($\times 10^9$/L)	173.0 ± 91.0	181.3 ± 97.9	8.3 ± 6.9	149.5 ± 74.0	157.6 ± 68.1	8.1 ± 5.9
Bilirubin (mg/dL)	10.8 ± 9.7	11.3 ± 8.3	0.5 ± 1.4	2.7 ± 2.8	3.5 ± 2.6	0.8 ± 0.2
Creatinine (mg/dL)	1.0 ± 1.0	1.4 ± 1.2	0.4 ± 0.2	0.9 ± 0.5	1.2 ± 1.2	0.3 ± 0.7

Table B2: Comparison of values computed in integrity check to reference article Table 1 (ALF)

Characteristic	Acute Liver Failure (ALF)					
	Publication: Non-APAP (n=50)	**DSIC: Non-APAP (n=50)**	**Diff. (n=0)**	**Publication: APAP (n=49)**	**DSIC: APAP (n=49)**	**Diff. (n=0)**
Age (years)	52.0 ± 22.2	50.8 ± 15.0	1.2 ± 7.2	38.0 ± 19.0	40.1 ± 13.8	2.1 ± 5.2
Gender (% female)	25 (50.0)	25 (50.0)	0 (0)	31 (63.3)	31 (63.3)	0 (0)
Race (% White)	40 (80.0)	40 (80.0)	0 (0)	46 (93.9)	46 (93.9)	0 (0)
INR	3.1 ± 2.6	0.4 ± 1.8	2.7 ± 0.8	3.4 ± 3.0	0.2 ± 1.3	3.2 ± 1.7
Platelets ($\times 10^9$/L)	118.0 ± 138.0	144.3 ± 111.0	26.3 ± 27.0	126.5 ± 87.8	141.4 ± 100.4	14.9 ± 12.6
Bilirubin (mg/dL)	7.4 ± 13.4	10.0 ± 8.2	2.6 ± 5.2	4.5 ± 4.4	4.9 ± 3.6	0.4 ± 0.8
Creatinine (mg/dL)	2.1 ± 2.0	2.7 ± 2.3	0.6 ± 0.3	1.7 ± 1.9	2.0 ± 1.6	0.3 ± 0.3

Attachment A: SAS Code

```
libname alfsg "X:\NIDDK\niddk-dr_studies1\ALFSG\private_orig_data\Final Public Use Data\ALFSG";

proc contents data=alfsg.form45;
run;

Proc contents data=alfsg.form00;
run;

proc contents data=alfsg.form16;
run;

proc contents data=alfsg.form02;
run;

proc contents data=alfsg.rotem;
run;

data rotem; set alfsg.rotem;
run;

proc sql;
select count(distinct subject_id) as distinct_id
from rotem;
quit;

/*********************/
/* DSIC for ALFSG */
/* Stravitz et al. */
/*********************/

*create temp datasets;
data dem; set alfsg.form00;
run;

data one; set alfsg.form45;
run;

data diag; set alfsg.form16;
run;

data age; set alfsg.form02;
keep age subject_id;
run;

data age1; set alfsg.form01;
```

```
keep age subject_id;
run;

*creating a singular age dataset;
data age2; set age age1;
run;

*merging to identify study subjects from paper;
proc sort data=dem;
by subject_id;
run;

proc sort data=one;
by subject_id;
run;

proc sort data=diag;
by subject_id;
run;

proc sort data=age2;
by subject_id;
run;

data two;
merge
dem (in=a)
one (in=b)
diag (in=c)
age (in=d);
by subject_id;
if b=1;
run;

*Table 1;
proc freq data=two;
tables f00alf f00ali;
run;

*removing xthe ten participants who converted from ALI to ALF over the course of the study;
data three; set two;
if f00alf = 1 AND f00ali = 1 then delete;
run;

proc freq data=three;
tables zVisitNm;
run;
```

```
data four; set three;
if f16q87 = 1 then apap = 1; if f16q87 ^= 1 then apap = 0;
run;

*APAP vs. Non-APAP;
proc freq data=four;
tables zVisitNm*apap;
run;

*ALI age;
data ali_age; merge
one (in=a)
age1 (in=b)
diag (in=c);
by subject_id;
if a=1;
run;

proc means data=ali_age mean std;
var Age;
where f16q87 = 1;
class zVisitNm;
run;

proc means data=ali_age mean std;
var Age;
where f16q87 ^= 1;
class zVisitNm;
run;

*ALF Age;
proc means data=four mean std;
var age;
class apap;
where zVisitNm = "ALF Admission";
run;

*gender;
proc sort data=four;
by zVisitNm;
run;

proc freq data=four;
tables Sex*apap/norow nopercent;
by zVisitNm;
run;

*race;
```

```
proc freq data=four;
tables White*apap/norow nopercent;
by zVisitNm;
run;

*INR;
data labs; set alfsg.form08;
run;

proc sort data=labs;
by subject_id zVisitNm;
run;

proc sort data=four;
by subject_id zVisitNm;
run;

data labs1; merge
four (in=a)
labs (in=b);
by subject_id zVisitNm;
if a=b;
run;

proc freq data=labs1;
tables zVisitNm;
run;

proc means data=labs1 mean std median q1 q3 min max;
var inr;
class apap;
where zVisitNm = "ALI Admission";
run;

proc means data=labs1 mean std median q1 q3 min max;
var inr;
class apap;
where zVisitNm = "ALF Admission";
run;

*platelets;
proc means data=labs1 mean std median q1 q3 min max;
var Platelet_Cnt;
class apap;
where zVisitNm = "ALI Admission";
run;

proc means data=labs1 mean std median q1 q3 min max;
```

```
var Platelet_Cnt;
class apap;
where zVisitNm = "ALF Admission";
run;

*Bilirubin;
proc means data=labs1 mean std median q1 q3;
var Bilirubin;
class apap;
where zVisitNm = "ALI Admission";
run;

proc means data=labs1 mean std median q1 q3;
var Bilirubin;
class apap;
where zVisitNm = "ALF Admission";
run;

*Creatinine;
proc means data=labs1 mean std median q1 q3;
var Creat;
class apap;
where zVisitNm = "ALI Admission";
run;

proc means data=labs1 mean std median q1 q3;
var Creat;
class apap;
where zVisitNm = "ALF Admission";
run;
```