Dataset Integrity Check for the Longitudinal Study of Genetic Causes of Intrahepatic Cholestasis (LOGIC) Boster

Prepared by NIDDK-CR
August 13, 2024

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

Cholestasis, a rare condition involving a reduction or obstruction of bile flow from the liver to the small intestine, can cause significant growth problems, liver complications, the need for liver transplantation, and death. The four rare genetic disorders, Alagille syndrome (ALGS), alpha-1 antitrypsin (a-1AT) deficiency, bile acid synthesis defects, and progressive familial intrahepatic cholestasis (PFIC), account for approximately 20% to 30% of all infant cases of cholestasis. Current knowledge concerning the etiology and outcomes of these diseases is limited. The Longitudinal Study of Genetic Causes of Intrahepatic Cholestasis (LOGIC) was established to investigate the natural history and progression of these four genetic disorders.

The main objective of the LOGIC Boster study on sarcopenia was to measure muscle mass in school-aged children with genetic intrahepatic cholestasis and assess relationships between sarcopenia, clinical variables, and outcomes.

3 Archived Datasets

All data files, as provided by the Data Coordinating Center (DCC), are located in the LOGIC folder in the data package. For this replication, variables were taken from the “sarcopenia.sas7bdat” dataset.

4 Statistical Methods

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

5 Results

For Table 1 in the publication [1], Demographic and clinical characteristics by disease, Table A lists the variables that were used in the replication, and Tables B1 and B2 compare the results calculated from the archived data files to the results published in Table 1. The results of the replication are within expected variation to the published results.

6 Conclusions

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

7 References

[1] Boster JM, Goodrich NP, Spino C, Loomes KM, Alonso EM, Kamath BM, Sokol RJ, Karpen S, Miethke A, Shneider BL, Molleston JP, Kohli R, Horslen SP, Rosenthal P, Valentino PL, Teckman JH, Hangartner TN, Sundaram SS. Sarcopenia is Associated With Osteopenia and Impaired Quality of Life in Children with Genetic Intrahepatic Cholestatic Liver Disease. Hepatology Communications, 7(11), e0293, October 2023. doi: https://doi.org/10.1097/HC9.0000000000000293

Table A: Variables used to replicate Table 1 – Demographic and clinical characteristics by disease

Table Variable	**dataset.variable**
Age	sarcopenia.ageatscanyears
Sex	sarcopenia.sex
Race	sarcopenia.race
Ethnicity	sarcopenia.ethnicity
Height z-score	sarcopenia.heightz
Weight z-score	sarcopenia.weightz
BMI	sarcopenia.bmi
Albumin	sarcopenia.albumingdl
INR	sarcopenia.inr
Total bilirubin	sarcopenia.bilitotalmgdl
Direct bilirubin	sarcopenia.bilidirectmgdl
GGT	sarcopenia.ggtpunitsl
Bile acids	sarcopenia.centralbileacid
AST	sarcopenia.astunitsl
Platelet count	sarcopenia.plateletscnt
APRI	sarcopenia.apri
Established CEPH prior to DEXA	sarcopenia.cephprescan
PedsQL Total (participant)	sarcopenia.pedsqltotalparticipant
PedsQL Total (parent)	sarcopenia.pedsqltotalparent
PedsQL Physical (participant)	sarcopenia.pedsqlphysparticipant
PedsQL Physical (parent)	sarcopenia.pedsqlphysparent

Table B1: Comparison of values computed in integrity check to reference article Table 1 (BASD and a1ATd)

Characteristic	Pub: BASD (n=12)	DSIC: BASD (n=12)	Diff. (n=0)	Pub: a1ATd (n=41)	DSIC: a1ATd (n=41)	Diff. (n=0)
Demographic						
Age, (years)	10.2 (5.4)	10.2 (5.4)	0 (0)	10.2 (4.6)	10.2 (4.6)	0 (0)
Sex, n (%)						
Female	3 (25.0)	3 (25.0)	0 (0)	12 (29.3)	12 (29.3)	0 (0)
Male	9 (75.0)	9 (75.0)	0 (0)	29 (70.7)	29 (70.7)	0 (0)
Race, n (%)						
Asian	1 (8.3)	1 (8.3)	0 (0)	-	-	-
Black or African American	1 (8.3)	1 (8.3)	0 (0)	-	-	-
White	9 (75.0)	9 (75.0)	0 (0)	40 (97.6)	40 (97.6)	0 (0)
Other or Multiracial	1 (8.3)	1 (8.3)	0 (0)	1 (2.4)	1 (2.4)	0 (0)
Unknown	-	-	-	-	-	-
Ethnicity, n (%)						
Hispanic/Latino	3 (25.0)	3 (25.0)	0 (0)	3 (7.3)	3 (7.3)	0 (0)
Non-Hispanic/Latino	9 (75.0)	9 (75.0)	0 (0)	37 (90.2)	37 (90.2)	0 (0)
Unknown	-	-	-	1 (2.4)	1 (2.4)	0 (0)
Clinical Characteristic						
Height z-score	-0.2 (0.7)	-0.2 (0.7)	0 (0)	0.5 (1.2)	0.5 (1.2)	0 (0)
Weight z-score	0.5 (1.1)	0.5 (1.1)	0 (0)	0.5 (0.9)	0.5 (0.9)	0 (0)
BMI (kg/m^2)	19.9 (5.1)	19.9 (5.1)	0 (0)	18.6 (3.9)	18.6 (3.9)	0 (0)
Albumin (g/dL)	4.6 (0.3)	4.6 (0.3)	0 (0)	4.4 (0.4)	4.4 (0.4)	0 (0)
INR	1.0 (0.1)	1.0 (0.1)	0 (0)	1.1 (0.1)	1.1 (0.1)	0 (0)
Total bilirubin (mg/dL)	0.4 (0.3)	0.4 (0.3)	0 (0)	0.6 (0.7)	0.6 (0.7)	0 (0)
Direct bilirubin (mg/dL)	-	-	-	(0.1)	(0.1)	(0)
GGT (IU/L)	23.6 (4.9)	23.6 (4.9)	0 (0)	57.5 (68.6)	57.5 (68.6)	0 (0)
Bile acids (µmol/L)	12.0 (5.7)	12.0 (5.7)	0 (0)	16.7 (19.6)	16.7 (19.6)	0 (0)
AST (IU/L)	54.1 (17.1)	54.1 (17.1)	0 (0)	62.3 (49.2)	62.3 (49.2)	0 (0)
Platelet count (10^3/mm)	269 (105)	269 (105)	0 (0)	238 (114)	238 (114)	0 (0)
APRI	0.6 (0.4)	0.6 (0.4)	0 (0)	(1.9)	(1.9)	0 (0)
Established CEPH prior to DEXA (n (%))	3 (25)	3 (25)	0 (0)	9 (22)	9 (22)	0 (0)
PedsQL Total (participant)	88.3 (6.8)	88.3 (6.8)	0 (0)	76.9 (16.3)	76.9 (16.3)	0 (0)
PedsQL Total (parent)	83.9 (13.4)	83.9 (13.4)	0 (0)	84.4 (16.1)	84.4 (16.1)	0 (0)
PedsQL Physical (participant)	89.3 (9.2)	89.3 (9.2)	0 (0)	80.6 (16.2)	80.6 (16.2)	0 (0)
PedsQL Physical (parent)	92.7 (5.2)	92.7 (5.2)	0 (0)	87.7 (18.4)	87.7 (18.4)	0 (0)

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

Characteristic	Pub: CIC (n=33)	DSIC: CIC (n=33)	Diff. (n=0)	Pub: ALGS (n=41)	DSIC: ALGS (n=41)	Diff. (n=0)
Demographic						
Age, (years)	11.2 (4.7)	11.2 (4.7)	0 (0)	9.8 (3.5)	9.8 (3.5)	0 (0)
Sex, n (%)						
Female	17 (51.5)	17 (51.5)	0 (0)	18 (43.9)	18 (43.9)	0 (0)
Male	16 (48.5)	16 (48.5)	0 (0)	25 (56.1)	23 (56.1)	2 (0)
Race, n (%)						
Asian	1 (3.0)	1 (3.0)	0 (0)	-	-	-
Black or African American	2 (6.1)	2 (6.1)	0 (0)	4 (9.8)	4 (9.8)	0 (0)
White	27 (81.8)	27 (81.8)	0 (0)	32 (78.0)	32 (78.0)	0 (0)
Other or Multiracial	1 (3.0)	1 (3.0)	0 (0)	3 (7.3)	3 (7.3)	0 (0)
Unknown	2 (6.1)	2 (6.1)	0 (0)	2 (4.9)	2 (4.9)	0 (0)
Ethnicity, n (%)						
Hispanic/Latino	7 (21.2)	7 (21.1)	0 (0)	3 (7.3)	3 (7.3)	0 (0)
Non-Hispanic/Latino	26 (78.8)	26 (78.8)	0 (0)	37 (90.2)	37 (90.2)	0 (0)
Unknown	-	-	-	1 (2.4)	1 (2.4)	0 (0)
Clinical Characteristic						
Height z-score	-1.2 (1.8)	-1.2 (1.8)	0 (0)	-1.7 (1.0)	-1.7 (1.0)	0 (0)
Weight z-score	-1.0 (1.6)	-1.0 (1.6)	0 (0)	-1.6 (1.3)	-1.6 (1.3)	0 (0)
BMI (kg/m^2)	17.7 (3.7)	17.7 (3.7)	0 (0)	16.1 (2.1)	16.1 (2.1)	0 (0)
Albumin (g/dL)	4.2 (0.5)	4.2 (0.5)	0 (0)	4.2 (0.6)	4.2 (0.6)	0 (0)
INR	1.1 (0.2)	1.1 (0.2)	0 (0)	1.1 (0.2)	1.1 (0.2)	0 (0)
Total bilirubin (mg/dL)	1.4 (1.8)	1.4 (1.8)	0 (0)	3.8 (5.2)	3.8 (5.2)	0 (0)
Direct bilirubin (mg/dL)	0.3 (0.6)	0.3 (0.6)	0 (0)	3.2 (3.4)	3.2 (3.4)	0 (0)
GGT (IU/L)	85.0 (163.4)	85.0 (163.4)	0 (0)	337.7 (302.6)	337.7 (302.6)	0 (0)
Bile acids (µmol/L)	78.0 (112.7)	78.0 (112.7)	0 (0)	131.7 (111.5)	131.7 (111.5)	0 (0)
AST (IU/L)	73.8 (43.1)	73.8 (43.1)	0 (0)	161.3 (114.3)	161.3 (114.3)	0 (0)
Platelet count (10^3/mm)	296 (139)	296 (139)	0 (0)	238 (109)	238 (109)	0 (0)
APRI	0.9 (1.1)	0.9 (1.1)	0 (0)	2.5 (2.4)	2.4 (2.5)	0.1 (0.1)
Established CEPH prior to DEXA (n (%))	5 (15)	5 (15)	0 (0)	18 (44)	18 (44)	0 (0)
PedsQL Total (participant)	72.4 (17.6)	72.4 (17.6)	0 (0)	73.7 (13.2)	73.6 (13.2)	0.1 (0)
PedsQL Total (parent)	77.4 (13.3)	77.4 (13.3)	0 (0)	71.5 (16.1)	71.5 (16.1)	0 (0)
PedsQL Physical (participant)	77.9 (19.5)	77.9 (19.5)	0 (0)	80.5 (15.3)	80.5 (15.3)	0 (0)
PedsQL Physical (parent)	79.7 (14.7)	79.7 (14.7)	0 (0)	71.7 (20.4)	71.7 (20.4)	0 (0)

Attachment A: SAS Code

```
libname sarc "X:\NIDDK\niddk-
dr_studies2\LOGIC\private_created_data\Sarcopenia\LOGIC_Boster_Sarcopenia_AN105\DataAndFiles";
libname library "X:\NIDDK\niddk-
dr_studies2\LOGIC\private_created_data\Sarcopenia\LOGIC_Boster_Sarcopenia_AN105\DataAndFiles";

*DSIC for LOGIC Boster Sarcopenia;

*disease categories;
proc freq data=sarc.sarcopenia;
tables Diagnosis;
run;

*Demographics;
*age;
proc means data=sarc.sarcopenia n median mean std;
var AgeAtScanYears;
class Diagnosis;
run;

proc freq data=sarc.sarcopenia;
tables (Sex Race Ethnicity)*Diagnosis/norow nopercent;
run;

*Clinical Characteristics;
proc means data=sarc.sarcopenia n mean std;
var HeightZ WeightZ bmi Albumingdl inr BiliTotalMgdl BiliDirectMgdl
GGTPUnitsL CentralBileAcid ASTUnitsL Plateletscnt apri
PedsQLTotalParticipant PedsQLTotalParent PedsQLPhysParticipant
PedsQLPhysParent;
class Diagnosis;
run;

proc freq data=sarc.sarcopenia;
tables CEPHPreScan*diagnosis/norow nopercent;
run;
```