

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

Prepared by NIDDK-CR
October 14, 2022

Contents

1 Standard Disclaimer	2
2 Study Background	2
3 Archived Datasets	2
4 Statistical Methods	2
5 Results	3
6 Conclusions	3
7 References	3
Table A: Variables used to replicate Table 1 – Demographics and clinical characteristics of children with inherited cholestatic liver disease	4
Table B1: Comparison of values computed in integrity check to reference article Table 1 (ALGS and PFIC)	6
Table B2: Comparison of values computed in integrity check to reference article Table 1 (A1AT and Total)	7
Attachment A: SAS Code	8

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 LOGIC Neurodevelopmental Outcomes study evaluated neurodevelopmental status among children with inherited cholestatic liver diseases with native liver and variables predictive of impairment.

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 “manuscript_data_03oct19.sas7bdat” dataset.

4 Statistical Methods

Analyses were performed to replicate results for the data in the publication by Leung 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 clinical characteristics of children with inherited cholestatic liver 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 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 data files to be distributed are a true copy of the study data.

7 References

[1] Leung DH, Sorensen LG, Ye W, Hawthorne K, Ng VL, Loomes KM, Fredericks EM, Alonso EM, Heubi JE, Horslen SP, Karpen SJ, Molleston JP, Rosenthal P, Sokol RJ, Squires RH, Wang KS, Kamath BM, Magee JC. Neurodevelopmental Outcomes in Children With Inherited Liver Disease and Native Liver. *Journal of Pediatric Gastroenterology and Nutrition*, 74(1), 96-103, January 2022. doi: <https://doi.org/10.1097/MPG.0000000000003337>

Table A: Variables used to replicate Table 1 – Demographics and clinical characteristics of children with inherited cholestatic liver disease

Table Variable	dataset.variable
Age at testing	manuscript_data_03oct19.dx manuscript_data_03oct19.age_test_y
Female	manuscript_data_03oct19.dx manuscript_data_03oct19.sex
Race	manuscript_data_03oct19.dx manuscript_data_03oct19.race
Hispanic	manuscript_data_03oct19.dx manuscript_data_03oct19.ethnicity
Parental education	manuscript_data_03oct19.dx manuscript_data_03oct19.college
Test type	manuscript_data_03oct19.dx manuscript_data_03oct19.test
Height-for-age z-score	manuscript_data_03oct19.dx manuscript_data_03oct19.haz
Weight-for-age z-score	manuscript_data_03oct19.dx manuscript_data_03oct19.waz
BMI z-score	manuscript_data_03oct19.dx manuscript_data_03oct19.bmiz
Pruritis	manuscript_data_03oct19.dx manuscript_data_03oct19.pruritis
History of xanthoma(s)	manuscript_data_03oct19.dx manuscript_data_03oct19.xanthoma
CEPH	manuscript_data_03oct19.dx manuscript_data_03oct19.ceph
Total bilirubin (mg/dL)	manuscript_data_03oct19.dx manuscript_data_03oct19.totalbilirubinmgdl
GGTP (U/L)	manuscript_data_03oct19.dx manuscript_data_03oct19.ggtpunitsl
AST (U/L)	manuscript_data_03oct19.dx manuscript_data_03oct19.astunitsl
ALT (U/L)	manuscript_data_03oct19.dx manuscript_data_03oct19.altunitsl
Alkaline phosphatase (U/L)	manuscript_data_03oct19.dx manuscript_data_03oct19.alkphosphunitsl
Platelet count (x 10 ³ /mm ³)	manuscript_data_03oct19.dx manuscript_data_03oct19.plateletsct
INR	manuscript_data_03oct19.dx manuscript_data_03oct19.inr
Albumin (g/dL)	manuscript_data_03oct19.dx manuscript_data_03oct19.albumingdl
BUN (mg/dL)	manuscript_data_03oct19.dx manuscript_data_03oct19.bunmgdl

Creatinine (mg/dL)	manuscript_data_03oct19.dx manuscript_data_03oct19.creatininemdgl
Hemoglobin (g/dL)	manuscript_data_03oct19.dx manuscript_data_03oct19.hgbgdl
APRI	manuscript_data_03oct19.dx manuscript_data_03oct19.apri
FIB-4	manuscript_data_03oct19.dx manuscript_data_03oct19.fib_4

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

Variable (Frequency (%) or mean (SD))	Publication: ALGS (n=70)	DSIC: ALGS (n=70)	Diff. (n=0)	Publication: PFIC (n=43)	DSIC: PFIC (n=43)	Diff. (n=0)
Age at testing (years)	8.7 (3.7)	8.7 (3.7)	0 (0)	8.5 (4.3)	8.5 (4.3)	0 (0)
Female	25 (35.7)	25 (35.7)	0 (0)	22 (51.2)	22 (51.2)	0 (0)
Race						
White	52 (77.6)	52 (77.6)	0 (0)	32 (74.4)	32 (74.4)	0 (0)
Black	6 (9.0)	6 (9.0)	0 (0)	5 (11.6)	5 (11.6)	0 (0)
Other	9 (13.4)	9 (13.4)	0 (0)	6 (14.0)	6 (14.0)	0 (0)
Hispanic	12 (17.1)	12 (17.1)	0 (0)	5 (11.9)	5 (11.9)	0 (0)
Parental education						
High school equivalent or less	14 (21.2)	14 (21.1)	0 (0)	12 (30.8)	12 (30.8)	0 (0)
Some college/trade school/more	52 (78.8)	52 (78.8)	0 (0)	27 (69.2)	27 (69.2)	0 (0)
Test type						
WPPSI-III (age 3-5)	19 (27.1)	19 (27.1)	0 (0)	17 (39.5)	17 (39.5)	0 (0)
WISC-IV (age 6-16)	51 (72.9)	51 (72.9)	0 (0)	26 (60.5)	26 (60.5)	0 (0)
Height-for-age z-score	-1.51 (1.18)	-1.51 (1.18)	0 (0)	-0.90 (1.24)	-0.90 (1.24)	0 (0)
Weight-for-age z-score	-1.34 (1.25)	-1.34 (1.25)	0 (0)	-0.47 (1.14)	-0.47 (1.14)	0 (0)
BMI z-score	-0.55 (1.29)	-0.58 (1.30)	0.03 (0.01)	0.23 (1.03)	0.23 (1.03)	0 (0)
Pruritis						
None/mild	42 (61.8)	42 (61.8)	0 (0)	32 (74.4)	32 (74.4)	0 (0)
Active/bleeding	26 (38.2)	26 (38.2)	0 (0)	11 (25.6)	11 (25.6)	0 (0)
History of xanthoma(s)	18 (26.1)	18 (26.1)	0 (0)	1 (2.3)	1 (2.3)	0 (0)
CEPH						
aCEPH	46 (65.7)	46 (65.7)	0 (0)	34 (79.1)	34 (79.1)	0 (0)
pCEPH	14 (20.0)	14 (20.0)	0 (0)	5 (11.6)	5 (11.6)	0 (0)
dCEPH	10 (14.3)	10 (14.3)	0 (0)	4 (9.3)	4 (9.3)	0 (0)
Total bilirubin (mg/dL)	1.6 (0.7, 3.4)	1.6 (0.7, 3.4)	0 (0)	0.7 (0.4, 1.9)	0.7 (0.4, 1.9)	0 (0)
GGTP (U/L)	353 (233, 869)	353 (233, 869)	0 (0)	19 (12, 48)	19 (12, 48)	0 (0)
AST (U/L)	153 (102, 192)	153 (102, 192)	0 (0)	60 (36, 113)	60 (36, 113)	0 (0)
ALT (U/L)	179 (119, 242)	179 (119, 242)	0 (0)	50 (32, 106)	50 (32, 106)	0 (0)
Alkaline phosphatase (U/L)	506 (369, 653)	506 (369, 653)	0 (0)	474 (331, 553)	474 (331, 553)	0 (0)
Platelet count (x 10 ³ /mm ³)	244 (171, 338)	244 (171, 338)	0 (0)	305 (206, 377)	305 (206, 377)	0 (0)
INR	1.0 (0.9, 1.1)	1.0 (0.9, 1.1)	0 (0)	1.0 (1.0, 1.1)	1.0 (1.0, 1.1)	0 (0)
Albumin (g/dL)	4.3 (4.0, 4.5)	4.3 (4.0, 4.5)	0 (0)	4.2 (3.9, 4.5)	4.2 (3.9, 4.5)	0 (0)
BUN (mg/dL)	16 (14, 19)	16 (14, 19)	0 (0)	12 (9, 14)	12 (9, 14)	0 (0)
Creatinine (mg/dL)	0.44 (0.37, 0.50)	0.44 (0.37, 0.50)	0 (0)	0.35 (0.30, 0.50)	0.35 (0.30, 0.50)	0 (0)
Hemoglobin (g/dL)	12.9 (12.0, 13.5)	12.9 (12.0, 13.5)	0 (0)	12.6 (11.8, 13.5)	12.6 (11.8, 13.5)	0 (0)
APRI	1.39 (0.98, 2.96)	1.39 (0.98, 2.96)	0 (0)	0.52 (0.33, 0.96)	0.52 (0.33, 0.96)	0 (0)
< 1	15 (25.9)	15 (25.9)	0 (0)	28 (77.8)	28 (77.8)	0 (0)
1-1.5	16 (27.6)	16 (27.6)	0 (0)	1 (2.8)	1 (2.8)	0 (0)
> 1.5	27 (46.6)	27 (46.6)	0 (0)	7 (19.4)	7 (19.4)	0 (0)
FIB-4	0.41 (0.21, 0.89)	0.41 (0.21, 0.89)	0 (0)	0.22 (0.14, 0.38)	0.22 (0.14, 0.38)	0 (0)

Table B2: Comparison of values computed in integrity check to reference article Table 1 (A1AT and Total)

Variable (Frequency (%) or mean (SD))	Publication: A1AT (n=102)	DSIC: A1AT (n=102)	Diff. (n=0)	Publication: Total (n=215)	DSIC: Total (n=215)	Diff. (n=0)
Age at testing (years)	7.7 (3.5)	7.7 (3.5)	0 (0)	8.2 (3.8)	8.2 (3.8)	0 (0)
Female	29 (28.4)	29 (28.4)	0 (0)	76 (35.3)	76 (35.3)	0 (0)
Race						
White	94 (92.2)	94 (92.2)	0 (0)	178 (84.0)	178 (84.0)	0 (0)
Black	0 (0.0)	0 (0.0)	0 (0)	11 (5.2)	11 (5.2)	0 (0)
Other	8 (7.8)	8 (7.8)	0 (0)	23 (10.8)	23 (10.8)	0 (0)
Hispanic	5 (5.0)	5 (5.0)	0 (0)	22 (10.4)	22 (10.4)	0 (0)
Parental education						
High school equivalent or less	11 (11.8)	11 (11.8)	0 (0)	37 (18.7)	37 (18.7)	0 (0)
Some college/trade school/more	82 (88.2)	82 (88.2)	0 (0)	161 (81.3)	161 (81.3)	0 (0)
Test type						
WPPSI-III (age 3-5)	42 (41.2)	42 (41.2)	0 (0)	78 (36.3)	78 (36.3)	0 (0)
WISC-IV (age 6-16)	60 (58.8)	60 (58.8)	0 (0)	137 (63.7)	137 (63.7)	0 (0)
Height-for-age z-score	0.45 (1.03)	0.45 (1.03)	0 (0)	-0.47 (1.43)	-0.47 (1.43)	0 (0)
Weight-for-age z-score	0.71 (1.05)	0.71 (1.05)	0 (0)	-0.18 (1.45)	-0.18 (1.45)	0 (0)
BMI z-score	0.70 (1.04)	0.70 (1.04)	0 (0)	0.20 (1.25)	0.18 (1.26)	0.02 (0.01)
Pruritis						
None/mild	98 (99.0)	98 (99.0)	0 (0)	172 (81.9)	172 (81.9)	0 (0)
Active/bleeding	1 (1.0)	1 (1.0)	0 (0)	38 (18.1)	38 (18.1)	0 (0)
History of xanthoma(s)	0 (0.0)	0 (0.0)	0 (0)	19 (8.9)	19 (8.9)	0 (0)
CEPH						
aCEPH	79 (77.5)	79 (77.5)	0 (0)	159 (74.0)	159 (74.0)	0 (0)
pCEPH	7 (6.9)	7 (6.9)	0 (0)	26 (12.1)	26 (12.1)	0 (0)
dCEPH	16 (15.7)	16 (15.7)	0 (0)	30 (14.0)	30 (14.0)	0 (0)
Total bilirubin (mg/dL)	0.4 (0.3, 0.7)	0.4 (0.3, 0.7)	0 (0)	0.6 (0.4, 1.6)	0.6 (0.4, 1.6)	0 (0)
GGTP (U/L)	31 (19, 91)	31 (19, 91)	0 (0)	55 (19, 275)	55 (19, 275)	0 (0)
AST (U/L)	62 (42, 93)	62 (42, 93)	0 (0)	77 (48, 147)	77 (48, 147)	0 (0)
ALT (U/L)	68 (43, 110)	68 (43, 110)	0 (0)	90 (46, 179)	90 (46, 179)	0 (0)
Alkaline phosphatase (U/L)	255 (211, 337)	255 (211, 337)	0 (0)	347 (236, 513)	347 (236, 513)	0 (0)
Platelet count (x 10 ³ /mm ³)	263 (201, 324)	263 (201, 324)	0 (0)	266 (189, 333)	266 (189, 333)	0 (0)
INR	1.1 (1.0, 1.1)	1.1 (1.0, 1.1)	0 (0)	1.0 (1.0, 1.1)	1.0 (1.0, 1.1)	0 (0)
Albumin (g/dL)	4.4 (4.1, 4.7)	4.4 (4.1, 4.7)	0 (0)	4.3 (4.0, 4.6)	4.3 (4.0, 4.6)	0 (0)
BUN (mg/dL)	12 (11, 15)	12 (11, 15)	0 (0)	13 (11, 16)	13 (11, 16)	0 (0)
Creatinine (mg/dL)	0.40 (0.33, 0.54)	0.40 (0.33, 0.54)	0 (0)	0.40 (0.30, 0.50)	0.40 (0.30, 0.50)	0 (0)
Hemoglobin (g/dL)	13.5 (13.0, 14.0)	13.5 (13.0, 14.0)	0 (0)	13.1 (12.3, 13.9)	13.1 (12.3, 13.9)	0 (0)
APRI	0.64 (0.37, 1.32)	0.64 (0.37, 1.32)	0 (0)	0.83 (0.41, 1.83)	0.83 (0.41, 1.83)	0 (0)
< 1	63 (72.4)	63 (72.4)	0 (0)	106 (58.6)	106 (58.6)	0 (0)
1-1.5	4 (4.6)	4 (4.6)	0 (0)	21 (11.6)	21 (11.6)	0 (0)
> 1.5	20 (23.0)	20 (23.0)	0 (0)	54 (29.8)	54 (29.8)	0 (0)
FIB-4	0.24 (0.13, 0.36)	0.24 (0.13, 0.36)	0 (0)	0.26 (0.14, 0.53)	0.26 (0.14, 0.53)	0 (0)

Attachment A: SAS Code

```
libname leung "X:\NIDDK\niddk-dr_studies2\LOGIC\private_orig_data\Leung (AN084) ALGS  
Neurodevelopment Data Submission";
```

```
/******  
/* DSIC for Leung Neurodev */  
/******
```

```
data dsic; set leung.manuscript_data_03oct19;  
run;
```

```
*keeping those participants that were tested;  
data dsic1; set dsic;  
where tested = 1;  
run;
```

```
*dx groups;  
proc freq data=dsic1;  
tables dx;  
run;
```

```
*age;  
proc sort data=dsic1;  
by dx;  
run;
```

```
proc means data=dsic1 n mean std;  
var age_test_y;  
class dx;  
run;
```

```
*female;  
proc freq data=dsic1;  
tables sex*dx/norow;  
run;
```

```
*race;  
proc freq data=dsic1;  
tables race*dx/norow;  
run;
```

```
*Ethnicity;  
proc freq data=dsic1;  
tables Ethnicity*dx/norow;  
run;
```

```
*Parental education;  
proc freq data=dsic1;  
tables college*dx/norow;  
run;
```

```
*Test type;  
proc freq data=dsic1;  
tables test*dx/norow;  
run;
```

```
*height z score;  
proc means data=dsic1 mean std;  
var haz;  
class dx;  
run;
```

```
*weight z score;  
proc means data=dsic1 mean std;  
var waz;  
class dx;  
run;
```

```
*BMI z score;  
proc means data=dsic1 mean std;  
var bmiz;  
class dx;  
run;
```

```
*Pruritis;  
proc freq data=dsic1;  
tables pruritus*dx/norow;  
run;
```

```
*History of xanthomas;  
proc freq data=dsic1;  
tables xanthoma*dx/norow;  
run;
```

```
*CEPH;  
proc freq data=dsic1;  
tables ceph*dx/norow;  
run;
```

```
*Total bilirubin;  
proc means data=dsic1 median q1 q3;  
var TotalBilirubinMgdI;  
class dx;  
run;
```

```
*GGTP;  
proc means data=dsic1 median q1 q3;  
var GGTPUnitsL;  
class dx;  
run;
```

```
*AST;  
proc means data=dsic1 median q1 q3;  
var ASTUnitsL;  
/*class dx; */  
run;
```

```
*ALT;  
proc means data=dsic1 median q1 q3;  
var ALTUnitsL;  
class dx;  
run;
```

```
*Alkaline phosphatase;  
proc means data=dsic1 median q1 q3;  
var AlkPhosphUnitsL;  
/*class dx; */  
run;
```

```
*Platelet;  
proc means data=dsic1 median q1 q3;  
var PlateletsCnt;  
class dx;  
run;
```

```
*INR;  
proc means data=dsic1 median q1 q3;  
var inr;  
class dx;  
run;
```

```
*Albumin;  
proc means data=dsic1 median q1 q3;  
var AlbuminGdl;  
class dx;  
run;
```

```
*BUN;  
proc means data=dsic1 median q1 q3;  
var BUNMgdl;  
/*class dx; */  
run;
```

```

*Creatinine;
proc means data=dsic1 median q1 q3;
var CreatinineMgdl;
class dx;
run;

*Hemoglobin;
proc means data=dsic1 median q1 q3;
var HgbGdl;
/*class dx; */
run;

*APRI;
proc means data=dsic1 median q1 q3;
var apri;
class dx;
run;

data dsic2; set dsic1;
apri_cat = .;
if apri < 1 AND apri > 0 then apri_cat = 1;
if apri >= 1 AND apri <= 1.5 then apri_cat = 2;
if apri > 1.5 then apri_cat = 3;
run;

proc freq data=dsic2;
tables apri_cat*dx/norow;
run;

*FIB-4;
proc means data=dsic1 median q1 q3;
var fib_4;
class dx;
run;

```