

# Dataset Integrity Check for LOGIC PedsQL Analysis Dataset

**Prepared by Michael Spriggs**

**IMS Inc.**

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

**August 7, 2018**

# Contents

1 Standard Disclaimer .....	2
2 Study Background .....	2
3 Archived Datasets .....	2
4 Statistical Methods .....	2
5 Results .....	2
6 Conclusions .....	3
7 References .....	3
Table A: Variables used to replicate Table 1: A Comparison of PedsQL™ 4.0 Generic Core Scale Scores in Patients with Alagille Syndrome and Healthy Children, Child Self and Parent Proxy Reports .....	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	5
Table C: Variables used to replicate Table 4: Demographic and Medical Characteristics of Patients with Alagille Syndrome, A1ATD, and CIC Completing Child Self-Report of HRQOL.....	6
Table D: Comparison of values computed in integrity check to reference article Table 4 values .....	7
Table E: Variables used to replicate Table 6: A Comparison of the PedsQL 4.0™ Generic Core Scale Scores in children with Alagille syndrome and Chronic Intrahepatic Cholestasis (CIC), child self and parent proxy reports.....	9
Table F: Comparison of values computed in integrity check to reference article Table 6 values .....	10
Attachment A: SAS Code.....	12

## 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) study is a longitudinal cohort study that was established by the Childhood Liver Disease Research and Education Network (ChiLDREN) to investigate the natural history and progression of these four genetic disorders.

## 3 Archived Datasets

All the SAS 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 pedsq<sub>l</sub>\_child\_dis and pedsq<sub>l</sub>\_parent\_dis datasets.

## 4 Statistical Methods

Analyses were performed to duplicate results for the data published by Kamath et al [1] in Journal of Pediatrics in 2015. To verify the integrity of the dataset, descriptive statistics were computed.

## 5 Results

For Table 1 in the publication [1], [A Comparison of PedsQL 4.0 Generic Core Scale Scores in Patients with Alagille Syndrome and Healthy Children, Child Self and Parent Proxy Reports](#), Table A lists the variables that were used in the replication and Table B compares the results calculated from the archived data files to the results published in Table 1. The results of the replication are identical when available. Per

the paper, the healthy population data is from Varni et al, 2003 and were not included in the data package.

For Table 4 in the publication [1], Demographic and Medical Characteristics of Patients with Alagille Syndrome, A1ATD, and CIC Completing Child Self-Report of HRQOL, Table C lists the variables that were used in the replication and Table D compares the results calculated from the archived data files to the results published in Table 4. The results of the replication found some similar results.

For Table 6 in the publication [1], A Comparison of the PedsQL 4.0™ Generic Core Scale Scores in children with Alagille syndrome and Chronic Intrahepatic Cholestasis (CIC), child self and parent proxy reports, Table E lists the variables that were used in the replication and Table F compares the results calculated from the archived data files to the results published in Table 6. The results of the replication are identical when available.

## **6 Conclusions**

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

## **7 References**

[1] Kamath et. al. Quality of Life and Its Determinants in a Multicenter Cohort of Children with Alagille Syndrome. J Pediatr. 2015 Aug;167(2):390-6.e3. doi: 10.1016/j.jpeds.2015.04.077. Epub 2015 Jun 6.

**Table A:** Variables used to replicate Table 1: A Comparison of PedsQL 4.0 Generic Core Scale Scores in Patients with Alagille Syndrome and Healthy Children, Child Self and Parent Proxy Reports

<b>Table Variable</b>	<b>dataset.variable</b>
Child Self-Report	
Total Score	pedsq1_child_dis.total
Physical Health	pedsq1_child_dis.physical
Psychosocial Health	pedsq1_child_dis.psychosocial
Emotional Functioning	pedsq1_child_dis.emotion
Social Functioning	pedsq1_child_dis.social
School Functioning	pedsq1_child_dis.school
Parent Proxy-Report	
Total Score	pedsq1_parent_dis.total
Physical Health	pedsq1_parent_dis.physical
Psychosocial Health	pedsq1_parent_dis.psychosocial
Emotional Functioning	pedsq1_parent_dis.emotion
Social Functioning	pedsq1_parent_dis.social
School Functioning	pedsq1_parent_dis.school

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

Variable	Alagille Syndrome Manuscript	Alagille Syndrome DSIC	Diff.
Child Self Report			
Total Score	70 69.86 (16.09)	70 69.86 (16.09)	0 0(0)
Physical Health	69 72.52 (18.28)	69 72.52 (18.28)	0 0(0)
Psychosocial Health	70 68.56 (17.63)	70 68.56 (17.63)	0 0(0)
Emotional Functioning	70 69.13 (20.61)	70 69.13 (20.61)	0 0(0)
Social Functioning	70 69.11 (23.84)	70 69.11 (23.84)	0 0(0)
School Functioning	67 67.28 (19.98)	67 67.28 (19.98)	0 0(0)
Parent Proxy-Report			
Total Score	98 71.63 (16.72)	98 71.63 (16.72)	0 0(0)
Physical Health	98 73.81 (23.73)	98 73.81 (23.73)	0 0(0)
Psychosocial Health	97 70.94 (15.36)	97 70.94 (15.36)	0 0(0)
Emotional Functioning	98 67.70 (18.11)	98 67.70 (18.11)	0 0(0)
Social Functioning	97 75.22 (18.61)	97 75.22 (18.61)	0 0(0)
School Functioning	81 68.11 (20.42)	81 68.11 (20.42)	0 0(0)

**Table C:** Variables used to replicate Table 4: Demographic and Medical Characteristics of Patients with Alagille Syndrome, A1ATD, and CIC Completing Child Self-Report of HRQOL

<b>Table Variable</b>	<b>dataset.variable</b>
Female	Demographics.Gender
White	Demographics.race
Age at survey (in years)	Demographics.age_at_test
Weight (Z score)adjusted for age	Demographics.Weight
Height (Z score) adjusted for age	Demographics.Height
Listed for Liver Transplant	Demographics.Tx_listing
Total Bilirubin (mg/dL)	Demographics.bt_n
ALT (U/L)	Demographics.ALT_n
Albumin (g/dL)	Demographics.albumin_n
GGTP (U/L)	Demographics.GGTP_n
INR	Demographics.inr_n
WBC (103/mm3)	Demographics.wbc_n
Platelet (103/mm3)	Demographics.platelets_n

**Table D:** Comparison of values computed in integrity check to reference article Table 4 values

Variable	ALGS (N=70) Manuscript	ALGS (N=70) DSIC	Diff.	a1-AT (N=95) Manuscript	a1-AT (N=94) DSIC	Diff.	PFIC (N=49) Manuscript	PFIC (N=49) DSIC	Diff.
Female	27 39%	27 39%	0, 0%	40 42%	40 43%	0 -1%	28 57%	28 57%	0
White	53/62 (missing=8) 85%	53/67 (missing =3) 79%	0/5 -6%	90/90 (missing=5) 100%	89/93 (missing=1) 95%	1/-3 5%	35/46 (missing=3) 76%	35/46 (missing=3) 76%	0
Age at survey (in years)	70, 9.4 (3.1)	70, 9.4 (3.1)	0, 0 (0)	95 9.5 (3.9)	94, 9.6(3.9)	1, -0.1 (0)	49 10.3 (3.9)	48 10.4 (3.9)	1, -0.1(0)
Weight (Z score) adjusted for age	70, 1.5 (1.2)	70, 1.5 (1.2)	0, 0(0)	93 0.5 (1.1)	92, 0.6(1.1)	1 -0.1(0)	49 -0.8 (1.6)	48 -0.9 (1.6)	1, -1(0)
Height (Z score) adjusted for age	70, 1.6 (1.2)	70, 1.6 (1.2)	0, 0(0)	93 0.1 (3.0)	92, 0.1(3.0)	1 0(0)	49 -1.2 (1.6)	48 -1.3 (1.6)	1, -1(0)
Listed for Liver Transplant	6 9%	6 9%	0	6 6%	6 6%	0	10 20%	10 20%	0
Total Bilirubin (mg/dL)	55, 4.4 (6.2)	55, 4.4 (6.2)	0, 0(0)	77 0.7 (0.8)	76, 0.7(0.8)	1 0(0)	42 2.9 (5.2)	42 2.9 (5.2)	0
ALT (U/L)	66, 160.3 (96.9)	66, 160.3 (96.9)	0, 0 (0)	94 65.1 (53.5)	93, 65.0 (53.8)	1 0.1 (-0.3)	45 75.0 (56.4)	45 75.0 (56.4)	0
Albumin (g/dL)	67, 4.2 (0.5)	67, 4.2 (0.5)	0, 0 (0)	92 4.4 (0.5)	91, 4.4 (0.5)	1 0(0)	45 4.3 (0.7)	45 4.3 (0.7)	0
GGTP (U/L)	55, 386.2 (351.6)	55, 386.2 (351.6)	0, 0 (0)	77 61.8 (85.0)	76, 62.4 (85.3)	1 -6 (-.3)	40 94.1 (169.1)	40 94.1 (169.1)	0
INR	60, 1.0 (0.2)	60, 1.0 (0.2)	0, 0 (0)	66 1.1 (0.1)	65, 1.1 (0.1)	1 0(0)	37 1.1 (0.2)	37 1.1 (0.2)	0
WBC (103/mm3)	62, 7.0 (3.6)	62, 7.0 (3.6)	0, 0 (0)	85 7.0 (5.5)	84, 7.0 (5.6)	1 0(-0.1)	38 5.8 (2.3)	38 5.8 (2.3)	0



Variable	ALGS (N=70) Manuscript	ALGS (N=70) DSIC	Diff.	a1-AT (N=95) Manuscript	a1-AT (N=94) DSIC	Diff.	PFIC (N=49) Manuscript	PFIC (N=49) DSIC	Diff.
Platelet (103/mm3)	61, 234.7 (98.9)	61, 234.7 (98.9)	0,0 (0)	86 249.3 (116.4)	85 249.6 (117.1)	1 -0.3 (-0.7)	38 256.8 (155.6)	38 256.8 (155.6)	0

**Table E:** Variables used to replicate Table 6: A Comparison of the PedsQL 4.0™ Generic Core Scale Scores in children with Alagille syndrome and Chronic Intrahepatic Cholestasis (CIC), child self and parent proxy reports

<b>Table Variable</b>	<b>dataset.variable</b>
Child Self-Report	
Total Score	pedsq_child_dis.total
Physical Health	pedsq_child_dis.physical
Psychosocial Health	pedsq_child_dis.psychosocial
Emotional Functioning	pedsq_child_dis.emotion
Social Functioning	pedsq_child_dis.social
School Functioning	pedsq_child_dis.school
Parent Proxy-Report	
Total Score	pedsq_parent_dis.total
Physical Health	pedsq_parent_dis.physical
Psychosocial Health	pedsq_parent_dis.psychosocial
Emotional Functioning	pedsq_parent_dis.emotion
Social Functioning	pedsq_parent_dis.social
School Functioning	pedsq_parent_dis.school

**Table F:** Comparison of values computed in integrity check to reference article Table 6 values

Variable	Alagille Syndrome Manuscript	Alagille Syndrome DSIC	Diff. (n=0)	CIC Manuscript	CIC DSIC	Diff. (n=0)
Child Self Report						
Total Score	70 69.86 (16.09) 70.55 (20.65)	70 69.86 (16.09) 70.55 (20.65)	0 0(0)	49 73.04 (15.80) 72.83 (21.74)	49 73.04 (15.80) 72.83 (21.74)	0 0(0)
Physical Health	69 72.52 (18.28) 75.00 (18.75)	69 72.52 (18.28) 75.00 (18.75)	0 0(0)	49 78.91 (16.06) 81.25 (15.63)	49 78.91 (16.06) 81.25 (15.63)	0 0(0)
Psychosocial Health	70 68.56 (17.63) 67.26 (26.67)	70 68.56 (17.63) 67.26 (26.67)	0 0(0)	49 69.92 (17.69) 71.67 (20.00)	49 69.92 (17.69) 71.67 (20.00)	0 0(0)
Emotional Functioning	70 69.13 (20.61) 70.00 (35.00)	70 69.13 (20.61) 70.00 (35.00)	0 0(0)	49 67.35 (21.56) 65.00 (25.00)	49 67.35 (21.56) 65.00 (25.00)	0 0(0)
Social Functioning	70 69.11 (23.84) 70.00 (35.00)	70 69.11 (23.84) 70.00 (35.00)	0 0(0)	49 76.26 (20.81) 80.00 (30.00)	49 76.26 (20.81) 80.00 (30.00)	0 0(0)
School Functioning	67 67.28 (19.98) 65.00 (35.00)	67 67.28 (19.98) 65.00 (35.00)	0 0(0)	48 65.94 (19.75) 62.50 (27.50)	48 65.94 (19.75) 62.50 (27.50)	0 0(0)

Variable	Alagille Syndrome Manuscript	Alagille Syndrome DSIC	Diff. (n=0)	CIC Manuscript	CIC DSIC	Diff. (n=0)
Parent Proxy-Report						
Total Score	98 71.63 (16.72) 73.81 (21.63)	98 71.63 (16.72) 73.81 (21.63)	0 0(0)	68 79.02 (14.22) 79.46 (26.09)	68 79.02 (14.22) 79.46 (26.09)	0 0(0)
Physical Health	98 73.81 (23.73) 82.81 (37.50)	98 73.81 (23.73) 82.81 (37.50)	0 0(0)	68 81.47 (18.67) 85.94 (23.44)	68 81.47 (18.67) 85.94 (23.44)	0 0(0)
Psychosocial Health	97 70.94 (15.36) 71.15 (20.00)	97 70.94 (15.36) 71.15 (20.00)	0 0(0)	68 77.50 (15.06) 77.50 (25.00)	68 77.50 (15.06) 77.50 (25.00)	0 0(0)
Emotional Functioning	98 67.70 (18.11) 67.50 (25.00)	98 67.70 (18.11) 67.50 (25.00)	0 0(0)	68 74.10 (17.11) 77.50 (27.50)	68 74.10 (17.11) 77.50 (27.50)	0 0(0)
Social Functioning	97 75.22 (18.61) 75.00 (30.00)	97 75.22 (18.61) 75.00 (30.00)	0 0(0)	68 84.78 (16.76) 90.00 (25.00)	68 84.78 (16.76) 90.00 (25.00)	0 0(0)
School Functioning	81 68.11 (20.42) 70.00 (25.00)	81 68.11 (20.42) 70.00 (25.00)	0 0(0)	55 70.52 (21.87) 70.00 (35.00)	55 70.52 (21.87) 70.00 (35.00)	0 0(0)

## Attachment A: SAS Code

```
/******  
STUDY NAME: LOGIC  
PROGRAM LOCATION: P:\niddk\ims_analysis\LOGIC\prog_initial_analysis\duplicate_paper.sas  
SOFTWARE: SAS v9.4 Unix  
PROGRAMMER: Michael Spriggs  
PROGRAM FUNCTION: Data redaction review for LOGIC data
```

```
*****/  
title 'LOGIC Redaction Review and DSIC';  
title2 '/prj/niddk/ims_analysis/LOGIC/prog_initial_analysis/summarize_logic_data.sas';  
%include "/prj/biolincc/ims_analysis/sas_macros/redaction_data_summary.sas";  
libname logicin "/prj/niddk/ims_analysis/LOGIC/private_orig_data/12_13_delivery/";
```

```
data child_univariate;  
    set logicin.child_univariate;
```

```
label CTXSTDID ="De-Identified Subject ID."  
    bt_n="Total Bilirubin (mg/dL)"  
    dis_grp="Disease Group"  
    grp="Eligibility group"  
    physical="QOL Physical"  
    psychosocial="QOL psychosocial"  
    total="QOL Total"  
    HAZ="Height z-scores"  
    WAZ="Weight z-scores"  
    ;
```

```
data demographics;  
    set logicin.demographics;
```

```
label CTXSTDID ="De-Identified Subject ID."  
    albumin_n="Albumin (g/dL)"  
    ALT_n="ALT (U/L)"  
    bt_n="Total Bilirubin (mg/dL)"  
    GGTP_n="GGTP (U/L)"  
    inr_n="INR"  
    platelets_n="Platelet (10^3/mm3)"  
    wbc_n="WBC (10^3/mm3)"  
    age_at_test="Age at QOL survey (years)"  
    race="race"
```

```

Gender="Gender"
tx_listing="Transplant Listing"
dis_grp="Disease Group"
grp="Eligibility group"
physical="QOL Physical"
total="QOL Total"
HAZ="Height z-scores"
WAZ="Weight z-scores"
;

data parent_univariate;
set logicin.parent_univariate;

label CTXSTDID ="De-Identified Subject ID."
inr_n="INR"
age_at_test="Age at QOL survey (years)"
cardiac_defect="Cardiac defect"
dis_grp="Disease Group"
grp="Eligibility group"
physical="QOL Physical"
psychosocial="QOL psychosocial"
total="QOL Total"
HAZ="Height z-scores"
WAZ="Weight z-scores"
;

data pedsq_child_dis;
set logicin.pedsq_child_dis(Rename=(from=form));

label CTXSTDID ="De-Identified Subject ID."
dis_grp="Disease Group"
grp="Eligibility group"
emotion="QOL emotion"
physical="QOL Physical"
psychosocial="QOL psychosocial"
school="QOL school"
social="QOL social"
total="QOL Total"
form="QOL form source"
;

proc means data=pedsq_child_dis n mean stddev maxdec=2;
var total physical psychosocial emotion social school;
where dis_grp="ALGS";
title3 "Table B: Child Self Report";

data pedsq_parent_dis;

```

```
set logicin.pedsq_parent_dis(Rename=(from=form));

label CTXSTDID ="De-Identified Subject ID."
      dis_grp="Disease Group"
      grp="Eligibility group"
      emotion="QOL emotion"
      physical="QOL Physical"
      psychosocial="QOL psychosocial"
      school="QOL school"
      social="QOL social"
      total="QOL Total"
      form="QOL form source"
      ;

proc means data=pedsq_parent_dis n mean stddev maxdec=2;
var total physical psychosocial emotion social school;
where dis_grp="ALGS";
title3 "Table B: Parent Proxy Report";
```