Dataset Integrity Check for Biliary Atresia Study in Infants and Children (BASIC) Native Liver

Prepared by NIDDK-CR
September 9, 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

The Biliary Atresia Study in Infants and Children (BASIC) is a prospective, observational study to collect pertinent clinical information and biospecimens to aid in the understanding of the disorder. Specific aims of the study include identifying the gene(s) implicated in the etiology of biliary atresia, identifying the polymorphisms that may influence disease progression, and characterizing the natural history of the older, non-transplanted participants with biliary atresia.

The BASIC Native Liver study sought to examine the medical status of children with biliary atresia (BA) with their native livers after hepatic portoenterostomy (HPE) surgery.

3 Archived Datasets

A full listing of the archived datasets included in the package can be found in the Roadmap document. All data files, as provided by the Data Coordinating Center (DCC), are located in the BASIC Native Liver folder in the data package. For this replication, variables were taken from the “analysis.sas7bdat” dataset.

4 Statistical Methods

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

5 Results

For Table 1 in the publication [1], Characteristics of Study Cohort of 219 subjects with BA, 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 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 BASIC Native Liver data files to be distributed are a true copy of the study data.

7 References

[1] Ng VL, Haber BH, Magee JC, Miethke A, Murray KF, Michail S, Karpen SJ, Kerkar N, Molleston JP, Romero R, Rosenthal P, Schwarz KB, Shneider BL, Turmelle YP, Alonso EM, Sherker AH, Sokol RJ. Medical Status of 219 Children With Biliary Atresia Surviving Long-term With Their Native Livers: Results From a North American Multicenter Consortium. The Journal of Pediatrics, 165(3), 539-546, September 2014. doi: https://doi.org/10.1016/j.jpeds.2014.05.038

Table A: Variables used to replicate Table 1 – Characteristics of Study Cohort of 219 subjects with BA

Table Variable	**dataset.variable**
Sex	analysis.ymha1b2
Race	analysis.ms_race
Age at Kasai	analysis.age_at_kasai
Current age at evaluation	analysis.age_at_now
Polysplenia syndrome	analysis.ymhs1c13a
Other congenital malformations present	analysis.cardiac_hx analysis.gi_hx
Biochemical characteristics at enrollment	analysis.yilbb1mg analysis.yilbb6ul analysis.yilbb7ul analysis.yilbb9gl analysis.yilbb13in analysis.yilbb33pl analysis.yilbb11ul

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

Characteristics	**Pub (n=219)**	**DSIC (n=219)**	**Diff. (n=0)**
Sex			
Male	95 (43.4)	95 (43.4)	0 (0)
Female	124 (56.6)	124 (56.6)	0 (0)
Race			
White	139 (63.5)	139 (63.5)	0 (0)
Black	31 (14.1)	31 (14.1)	0 (0)
Asian	23 (10.5)	23 (10.5)	0 (0)
Other	26 (11.9)	26 (11.9)	0 (0)
Age at Kasai (days)			
N	183	183	0 (0)
Mean (SD)	56.4 (23.19)	56.4 (23.19)	0 (0)
Median (Range)	56 (7-125)	56 (7-125)	0 (0)
Current age at evaluation (years)			
N	219	219	0
Mean (SD)	10.5 (3.98)	13.5 (4.27)	3 (0.29)
Median (Range)	9.7 (5.1-17.9)	12.8 (5.3-22.8)	3.1 (0.2-4.9)
Polysplenia syndrome	5 (2.3)	5 (2.3)	0 (0)
Other congenital malformations present, n (%)			
Cardiac	23 (10.5)	23 (10.5)	0 (0)
Gastrointestinal	25 (11.4)	25 (11.4)	0 (0)
Biochemical characteristics at enrollment, median (IQR)			
TB, mg/dL	0.8 (0.4-1.2)	0.8 (0.4-1.2)	0 (0)
AST, IU/L	55 (38-99)	55 (38-99)	0 (0)
ALT, IU/L	54 (30-98)	54 (30-98)	0 (0)
Alb, g/dL	4.2 (3.8-4.5)	4.2 (3.8-4.5)	0 (0)
INR	1.1 (1.0-1.2)	1.1 (1.0-1.2)	0 (0)
Platelet count x 10^9/L	138 (85-209)	138 (85-209)	0 (0)
GGT, IU/L	67 (29-161)	67 (29-161)	0 (0)

Attachment A: SAS Code

```
libname basic "X:\NIDDK\niddk-dr_studies1\BASIC\private_orig_data\ChiLDReN BASIC Native Liver Dataset 201711";

/**************************/
/* BASIC Native Liver DSIC */
/* Ng et al. */
/**************************/

proc contents data=basic.analysis;
run;

*Categorical demographics;
proc freq data=basic.analysis;
tables YMHA1B2 ms_Race YMHS1C13A cardiac_hx GI_hx/missing;
run;

*Numerical demographics;
proc means data=basic.analysis n mean std median min max;
var age_at_kasai age_at_now age_at_consent Age_at_baseline ;
run;

*Labs;
proc means data=basic.analysis n median q1 q3;
var yilbb1mg yilbb6ul yilbb7ul yilbb9gl yilbb13in yilbb33pl yilbb11ul;
run;
```