Dataset Integrity Check for PROBE Total Bilirubin Analysis Dataset

Prepared by Dominick Parisi
IMS Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD 20705
August 7, 2018

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

PROBE is a multi-center project to establish a prospective database of clinical information and a repository of blood and tissue samples from children with diagnoses of neonatal liver diseases, such as biliary atresia and neonatal hepatitis, in order to perform research in these liver problems. Children were screened and enrolled at presentation at the participating pediatric liver sites. Subjects diagnosed with biliary atresia were followed intensively for the first year, at 18 months of age, and then annually up to 15 years of age. Other subjects diagnosed with cholestasis were followed on the same schedule; if there was complete (clinical and biochemical) resolution of their underlying liver disease off all therapy, there was one follow up visit within one year (preferably scheduled at the time of the next planned follow up visit or at 12 months of age, whichever was later) for data collection and to obtain blood samples. The development of a serum and tissue bank of specimens from children with various neonatal cholestatic disorders are used for future investigations into the etiology and pathogenesis of hepatobiliary injury in the infant.

3 Archived Datasets

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

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Benjamin L. Shneider et al [1] in The Journal of Pediatrics 2016. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], BA outcomes by 2 years of age, 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 similar to the published results.

6 Conclusions

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

7 References

[1] Benjamin L. Shneider, John C. Magee, MD, Saul J. Karpen, MD, PhD, Elizabeth B. Rand, MD, Michael R. Narkewicz, MD, Lee M. Bass, MD, Kathleen Schwarz, MD, Peter F. Whitington, MD, Jorge A. Bezerra, MD, Nanda Kerkar, MD, Barbara Haber, MD, Philip Rosenthal, MD, Yumirle P. Turmelle, MD, Jean P. Molleston, MD, Karen F. Murray, MD, Vicky L. Ng, MD, Kasper S. Wang, MD, Rene Romero, MD, Robert H. Squires, MD, Ronen Arnon, MD, Averell H. Sherker, MD, Jeffrey Moore, MS, Wen Ye, PhD, Ronald J. Sokol, MD on behalf of the show Childhood Liver Disease Research Network (ChiLDReN). Total Serum Bilirubin within 3 Months of Hepatoportoenterostomy Predicts Short-Term Outcomes in Biliary Atresia. The Journal of Pediatrics, March 2016, Vol. 170, Pages 211-217.e2.

Table A: Variables used to replicate Table 1: BA outcomes at 2 years of age

Table Variable	**dataset.variable**
Weight z-score	ba_outcomes_transfer_file.waz_event
Height z-score	ba_outcomes_transfer_file.haz_event
Mid-arm circumference z-score	ba_outcomes_transfer_file.armcirc_z_event
Hypoalbuminemia (albumin < 3.0 g/dL)	ba_outcomes_transfer_file.albumin_event
Coagulopathy (INR > 1.5)	ba_outcomes_transfer_file.inr_event
Hypnatremia (Na <130 mEq/L)	ba_outcomes_transfer_file.sodium_event
Thrombocytopenia (Platelets <150,000 uL)	ba_outcomes_transfer_file.platelets_event
Splenomegaly	ba_outcomes_transfer_file.spleen_event
Ascites	ba_outcomes_transfer_file.ascites_event
Variceal hemorrhage	ba_outcomes_transfer_file.varices_event
Liver transplant	ba_outcomes_transfer_file. time_from_birth_to_transplant
Death	ba_outcomes_transfer_file.final_status

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

Variable	PROBE Manuscript All BA (n=137) N (%)	PROBE DISC All BA (n=137) N (%)	Diff. (n=0)	PROBE Manuscript TB <2 mg/dL (n=68) N (%)	PROBE DSIC TB <2 mg/dL (n=68) N (%)	Diff. (n=0)
Weight z-score (< -2.5)	28 (20.4%)	28 (20.4%)	0 (0)	8 (11.8%)	8 (11.8%)	0 (0)
Height z-score (< -2.5)	26 (18.9%)	26 (19.0%)	0 (0)	10 (14.7%)	10 (14.7%)	0 (0)
Mid-arm circumference z-score (< -2.5)	51 (37.2%)	51 (37.2%)	0 (0)	24 (35.3%)	24 (35.3%)	0 (0)
Hypoalbuminemia (albumin < 3.0 g/dL)	46 (33.6%)	46 (33.6%)	0 (0)	9 (13.2%)	9 (13.2%)	0 (0)
Coagulopathy (INR > 1.5)	26 (19.0%)	26 (18.0%)	0 (0)	3 (4.4%)	3 (4.4%)	0 (0)
Hypnatremia (Na <130 mEq/L)	5 (3.7%)	5 (3.7%)	0 (0)	1 (1.5%)	1 (1.5%)	0 (0)
Thrombocytopenia (Platelets <150,000 uL)	65 (47.5%)	65 (47.5%)	0 (0)	32 (47.1%)	32 (47.1%)	0 (0)
Splenomegaly	86 (62.8%)	86 (62.8%)	0 (0)	48 (70.6%)	48 (70.6%)	0 (0)
Ascites	52 (38.0%)	52 (38.0%)	0 (0)	12 (17.7%)	12 (17.7%)	0 (0)
Variceal hemorrhage	9 (6.6%)	9 (6.6%)	0 (0)	5 (7.4%)	7 (10.3%)	2 (2.9)
Liver transplant	57 (41.6%)	58 (42.3%)	1 (0.7)	10 (14.7%)	10 (14.7%)	0 (0)
Death	9 (6.6%)	9 (6.6%)	0 (0)	2 (2.9%)	2 (2.9%)	0 (0)
Death or Liver transplant	66 (48.2%)	67 (48.9%)	1 (0.7)	12 (17.7%)	12 (17.7%)	0 (0)

Variable	PROBE Manuscript TB >=2 mg/dL (n=69) N (%)	PROBE DISC TB >=2 mg/dL (n=69) N (%)	Diff. (n=0)
Weight z-score (< -2.5)	20 (29.0%)	20 (29.0%)	0 (0)
Height z-score (< -2.5)	16 (23.2%)	16 (23.2%)	0 (0)
Mid-arm circumference z-score (< -2.5)	27 (39.1%)	27 (39.1%)	0 (0)
Hypoalbuminemia (albumin < 3.0 g/dL)	37 (53.6%)	37 (53.6%)	0 (0)
Coagulopathy (INR > 1.5)	23 (33.3%)	23 (33.3%)	0 (0)
Hypnatremia (Na <130 mEq/L)	4 (5.8%)	4 (5.8%)	0 (0)
Thrombocytopenia (Platelets <150,000 uL)	33 (47.8%)	33 (47.8%)	0 (0)
Splenomegaly	38 (55.1%)	38 (55.1%)	0 (0)
Ascites	40 (58.0%)	40 (58.0%)	0 (0)
Variceal hemorrhage	4 (5.8%)	2 (2.9%)	2 (2.9)
Liver transplant	47 (68.1%)	48 (69.6%)	1 (1.5)
Death	7 (10.1%)	7 (10.1%)	0 (0)
Death or Liver transplant	54 (78.3%)	55 (79.7%)	1 (1.4)

Attachment A: SAS Code

```
LIBNAME SASDATA '/prj/niddk/ims_analysis/PROBE/private_orig_data/ChiLDReN_Shneider_PROBE_TotalBilirubin Dataset/SAS Transfer File';

/*******************/
/* Import datasets */
/*******************/
DATA BA_OUTCOMES;
  SET SASDATA.BA_OUTCOMES_TRANSFER_FILE;

  IF      .<TOTAL_BILI_3<2  THEN TB_STATUS=1;
  ELSE IF   TOTAL_BILI_3>=2 THEN TB_STATUS=2;

  ** Jeff Moore response from 2/14/18;
  if time_from_birth_to_transplant ne . then transplant = 1;

RUN;

/************/
/* Table 1 */
/************/
TITLE2 'Table 1';
PROC FREQ DATA=BA_OUTCOMES;
  TABLE WAZ_EVENT
        HAZ_EVENT
        ARMCIRC_Z_EVENT
        ALBUMIN_EVENT
        INR_EVENT
        SODIUM_EVENT
        PLATELETS_EVENT
        SPLEEN_EVENT
        ASCITES_EVENT
        VARICES_EVENT
        transplant
        transplant*FINAL_STATUS
         /LIST MISSING;
RUN;

PROC FREQ DATA=BA_OUTCOMES;
  TABLE TB_STATUS * (WAZ_EVENT
                     HAZ_EVENT
                     ARMCIRC_Z_EVENT
                     ALBUMIN_EVENT
                     INR_EVENT
                     SODIUM_EVENT
                     PLATELETS_EVENT
                     SPLEEN_EVENT
                     ASCITES_EVENT
                     VARICES_EVENT
```

```
                     transplant
                    FINAL_STATUS) /MISSING NOPERCENT NOCOL;
  table tb_status*transplant*final_status/list missing nopercent nocol;
RUN;
```