Dataset Integrity Check for A Prospective Database of Infants with Cholestasis / A Randomized, Double-Blinded, Placebo-Controlled Trial of Corticosteroid Therapy Following Portoenterostomy in Infants With Biliary Atresia (PROBE/START) Alonso

Prepared by NIDDK-CR
May 3, 2024

Contents

1 Standard Disclaimer 2

2 Study Background 2

3 Archived Datasets 2

4 Statistical Methods 3

5 Results 3

6 Conclusions 3

7 References 3

Table A: Variables used to replicate Table 3 – Participant Characteristics 4

Table B: Comparison of values computed in integrity check to reference article Table 3 5

Attachment A: SAS Code 6

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

Two studies from the Childhood Liver Disease Research Network (ChiLDReN), A Prospective Database of Infants with Cholestasis (PROBE) and A Randomized, Double-Blinded, Placebo-Controlled Trial of Corticosteroid Therapy Following Portoenterostomy in Infants with Biliary Atresia (START), were used in assessing the impact of corticosteroid therapy on the growth of participants.

PROBE

The PROBE study was a multi-center project to establish a prospective database of clinical information and a repository of blood and tissue specimens from children with diagnoses of neonatal liver diseases, such as biliary atresia and neonatal hepatitis, in order to perform research on these liver problems. Children were screened and enrolled at presentation at the participating pediatric liver sites.

START

The START study participants were recruited from the PROBE study and randomized into either the corticosteroid or placebo group within 72 hours after the portoenterostomy procedure. Participants were given their assigned treatments daily over the course of 13 weeks. After the treatment period, participants underwent follow-up testing and assessments until age 24 months.

3 Archived Datasets

A full listing of 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 PROBE and START folders in the data packages. For this replication, variables were taken from the "start_growth_25oct18.sas7bdat" dataset.

4 Statistical Methods

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

5 Results

For Table 3 in the publication [1], Participant Characteristics, 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 3. The results of the replication are an exact match to the published results.

6 Conclusions

The NIDDK Central Repository is confident that the PROBE/START Alonso data files to be distributed are a true copy of the study data.

7 References

[1] Alonso EM, Ye W, Hawthorne K, Venkat V, Loomes KM, Mack CL, Hertel PM, Karpen SJ, Kerkar N, Molleston JP, Murray KF, Romero R, Rosenthal P, Schwarz KB, Shneider BL, Suchy FJ, Turmelle YP, Wang KS, Sherker AH, Sokol RJ, Bezerra JA, Magee JC. Impact of Steroid Therapy on Early Growth in Infants with Biliary Atresia: The Multicenter Steroids in Biliary Atresia Randomized Trial. The Journal of Pediatrics, 202, 179-185, November 2018. doi: https://doi.org/10.1016/j.jpeds.2018.07.002

Table A: Variables used to replicate Table 3 – Participant Characteristics

Table Variable	**dataset.variable**
Age at HPE (months)	start_growth_25oct18.age_hpe_d start_growth_25oct18.steroid start_growth_25oct18.visit
Age at HPE	start_growth_25oct18.age_hpe_d start_growth_25oct18.steroid start_growth_25oct18.visit
Weight at HPE (kg)	start_growth_25oct18.weightkg_hpe start_growth_25oct18.steroid start_growth_25oct18.visit
Weight z-score at HPE	start_growth_25oct18.waz_hpe start_growth_25oct18.steroid start_growth_25oct18.visit
Length z-score at HPE	start_growth_25oct18.haz_hpe start_growth_25oct18.steroid start_growth_25oct18.visit
Sex	start_growth_25oct18.sex start_growth_25oct18.steroid start_growth_25oct18.visit
Race	start_growth_25oct18.race start_growth_25oct18.steroid start_growth_25oct18.visit
Ethnicity	start_growth_25oct18.ethnicity start_growth_25oct18.steroid start_growth_25oct18.visit
BASM	start_growth_25oct18.basm start_growth_25oct18.steroid start_growth_25oct18.visit

Table B: Comparison of values computed in integrity check to reference article Table 3

Characteristic	Pub: Steroids (n=70)	DSIC: Steroids (n=70)	Diff. (n=0)	Pub: Placebo (n=70)	DSIC: Placebo (n=70)	Diff. (n=0)
Age at HPE (months)						
Mean (SD)	2.26 (0.93)	2.26 (0.93)	0 (0)	2.27 (0.84)	2.27 (0.84)	0 (0)
Min, max	0.85, 5.55	0.85, 5.55	0, 0	0.95, 5.26	0.95, 5.26	0, 0
Age at HPE						
≤ 30 days	2 (3%)	2 (3%)	0 (0)	2 (3%)	2 (3%)	0 (0)
> 30 and ≤ 45 days	17 (24%)	17 (24%)	0 (0)	8 (11%)	8 (11%)	0 (0)
> 45 and ≤ 60 days	10 (14%)	10 (14%)	0 (0)	18 (26%)	18 (26%)	0 (0)
> 60 and ≤ 90 days	30 (43%)	30 (43%)	0 (0)	29 (41%)	29 (41%)	0 (0)
> 90 and ≤ 120 days	7 (10%)	7 (10%)	0 (0)	11 (16%)	11 (16%)	0 (0)
> 120 days	4 (6%)	4 (6%)	0 (0)	2 (3%)	2 (3%)	0 (0)
Weight at HPE (kg)						
Mean (SD)	4.67 (0.97)	4.67 (0.97)	0 (0)	4.64 (0.82)	4.64 (0.82)	0 (0)
Min, max	2.6, 7.37	2.6, 7.37	0, 0	2.5, 6.8	2.5, 6.8	0, 0
Weight z-score at HPE						
Mean (SD)	-0.89 (1.09)	-0.89 (1.09)	0 (0)	-0.87 (1.02)	-0.87 (1.02)	0 (0)
Min, max	-3.56, 2.45	-3.56, 2.45	0, 0	-3.67, 1.6	-3.67, 1.6	0, 0
Length z-score at HPE						
Mean (SD)	-0.81 (1.36)	-0.81 (1.36)	0 (0)	-0.62 (1.13)	-0.62 (1.13)	0 (0)
Min, max	-4.94, 1.96	-4.94, 1.96	0, 0	-3.27, 1.67	-3.27, 1.67	0, 0
Sex						
Female	32 (46%)	32 (46%)	0 (0)	40 (57%)	40 (57%)	0 (0)
Male	38 (54%)	38 (54%)	0 (0)	30 (43%)	30 (43%)	0 (0)
Race						
White/Caucasian	46 (66%)	46 (66%)	0 (0)	44 (63%)	44 (63%)	0 (0)
Black/African American	8 (11%)	8 (11%)	0 (0)	11 (16%)	11 (16%)	0 (0)
Asian	6 (9%)	6 (9%)	0 (0)	2 (3%)	2 (3%)	0 (0)
American Indian/Alaskan Native	0 (0%)	0 (0%)	0 (0)	1 (1%)	1 (1%)	0 (0)
Other	7 (10%)	7 (10%)	0 (0)	9 (13%)	9 (13%)	0 (0)
Refused/Not reported	3 (4%)	3 (4%)	0 (0)	3 (4%)	3 (4%)	0 (0)
Ethnicity						
Hispanic	14 (20%)	14 (20%)	0 (0)	22 (31%)	22 (31%)	0 (0)
Non-Hispanic	55 (79%)	55 (79%)	0 (0)	48 (69%)	48 (69%)	0 (0)
Refused/Not reported	1 (1%)	1 (1%)	0 (0)	0 (0%)	0 (0)	0 (0)
BASM						
No	68 (97%)	68 (97%)	0 (0)	67 (96%)	67 (96%)	0 (0)
Yes	2 (3%)	2 (3%)	0 (0)	3 (4%)	3 (4%)	0 (0)

Attachment A: SAS Code

```
libname growth "X:\NIDDK\niddk-
dr_studies6\PROBE\private_created_data\Alonso\Alonso_START_growth";
libname library "X:\NIDDK\niddk-
dr_studies6\PROBE\private_created_data\Alonso\Alonso_START_growth";

/**********************************/
/* START/PROBE Growth Manuscript */
/* DSIC */
/**********************************/

*Age at HPE in months;
data one; set growth.start_growth_25oct18;
age_hpe_mos = (age_hpe_d/30.4);
run;

proc means data=one n mean std min max;
var age_hpe_mos;
class steroid;
where visit = 0;
run;

*Age at HPE categorical;
data two; set one;
if age_hpe_d <= 30 then age_cat = 1;
if age_hpe_d > 30 AND age_hpe_d <=45 then age_cat = 2;
if age_hpe_d > 45 AND age_hpe_d <=60 then age_cat = 3;
if age_hpe_d > 60 AND age_hpe_d <=90 then age_cat = 4;
if age_hpe_d > 90 AND age_hpe_d <=120 then age_cat= 5;
if age_hpe_d > 120 then age_cat = 6;
run;

proc freq data=two;
tables age_cat*steroid/norow nopercent;
where visit = 0;
run;

*Weight at HPE;
proc means data=two n mean std min max;
var weightkg_hpe;
class steroid;
```

```
where visit = 0;
run;

*Weight and length z-score at HPE;
proc means data=two n mean std min max;
var waz_hpe haz_hpe ;
class steroid;
where visit = 0;
run;

*Sex, race, ethnicity, BASM;
proc freq data=two;
tables (sex race ethnicity basm)*steroid/norow nopercent;
where visit = 0;
run;
```