Dataset Integrity Check for A Randomized, Double-Blinded, Placebo-Controlled Trial of Corticosteroid Therapy Following Portoenterostomy in Infants With Biliary Atresia (START) Venkat

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 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 START folder in the data package. For this replication, variables were taken from the "zenkat_archive.sas7bdat" dataset.

4 Statistical Methods

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

5 Results

For the Results section in the publication [1], 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 section. The results of the replication are within expected variation to the published results.

6 Conclusions

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

7 References

[1] Venkat VL, Shneider BL, Magee JC, Turmelle Y, Arnon R, Bezerra JA, Hertel PM, Karpen SJ, Kerkar N, Loomes KM, Molleston J, Murray KF, Ng VL, Raghunathan T, Rosenthal P, Schwartz K, Sherker AH, Sokol RJ, Teckman J, Wang K, Whitington PF, Heubi JE. Total Serum Bilirubin Predicts Fat-soluble Vitamin Deficiency Better than Serum Bile Acids in Infants with Biliary Atresia. Journal of Pediatric Gastroenterology and Nutrition, 59(6), 702-707, December 2014. doi: https://doi.org/10.1097/MPG.0000000000000547

Table A: Variables used to replicate demographics from the Results section

Table Variable	dataset.variable
Infants with biliary atresia	zenkat_archive.ctxvisno
Information at follow up after HPE entry	zenkat_archive.ctxvisno
Sex	zenkat_archive.dmab02mf
Ethnicity	zenkat_archive.dmab03hl
Race	zenkat_archive.race

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

Characteristic	Publication (n=92)	DSIC (n=92)	Diff. (n=0)
Infants with biliary atresia	92	92	0
Information at follow up after HPE entry	92	92	0
1 month	86	86	0
3 months	83	83	0
6 months	61	61	0
Sex			
Male	47%	47%	0
Ethnicity			
Hispanic	23%	24%	1
Race			
White	60%	62%	2
Black	16%	17%	1
Other (Asian, Native American, Unknown)	24%	21%	3

Attachment A: SAS Code

```
libname zenkat "X:\NIDDK\niddk-dr_studies6\START\private_created_data\START
Venkat\Venkat_START_VitaminDeficiency";

proc freq data=zenkat.zenkat_archive;
run;

/*********************/
/* DSIC for START */
/* Venkat */
/*********************/

*infants with BA, number at each visit, sex, ethnicity, and race;
proc freq data=zenkat.zenkat_archive;
tables ctxvisno dmab02mf dmab03hl race/missing;
run;
```