Dataset Integrity Check for the Clinical Trial of Entecavir/Pegylated Interferon in Immune Tolerant Children with Chronic Hepatitis B Virus (HBV) Infection (HBRN Immune Tolerant P)

Prepared by NIDDK-CR
October 20, 2023

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 Hepatitis B Research Network (HBRN) was a multicenter network to investigate the etiology and progression of the disease and to test the safety and efficacy of treatment approaches. The Clinical Trial of Entecavir/Pegylated Interferon in Immune Tolerant Children with Chronic Hepatitis B Virus (HBV) Infection (HBRN Immune Tolerant P) study was designed to evaluate the safety and efficacy of treatment using a combination of drugs (entecavir and pegylated interferon) in children ages 3 to < 18 years old with immunotolerant chronic hepatitis B.

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 HBRN Immune Tolerant P folder in the data package. For this replication, variables were taken from the “itp_long.sas7bdat” dataset.

4 Statistical Methods

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

5 Results

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

6 Conclusions

The NIDDK Central Repository is confident that the HBRN Immune Tolerant P data files to be distributed are a true copy of the study data.

7 References

[1] Rosenthal P, Ling SC, Belle SH, Murray KF, Rodriguez-Baez N, Schwarzenberg SJ, Teckman J, Lin HS, Schwarz KB. Combination of Entecavir/Peginterferon Alfa-2a in Children With Hepatitis B e Antigen-Positive Immune Tolerant Chronic Hepatitis B Virus Infection. Hepatology, 69(6), 2326-2337, June 2019. doi: https://doi.org/10.1002/hep.30312

Table A: Variables used to replicate Table 1 – Baseline Characteristics

Table Variable	**dataset.variable**
Age	itp_long.age_visit
Sex	itp_long.sex
BMI	itp_long.bmi
Race	itp_long.race
Born in the United States/Canada	itp_long.born_status
HBV genotype	itp_long.cdcgen_cat
HBV DNA	itp_long.hbvdna_c
Quantitative HBeAg	itp_long.equant
Quantitative HBsAg	itp_long.squant
ALT	itp_long.alt
ALT $\leqslant$ new ULN	itp_long.alt itp_long.alt_norm_new
AST	itp_long.ast
Total bilirubin	itp_long.tbili
Albumin	itp_long.alb
Platelets	itp_long.plat
White blood cells	itp_long.wbc

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

Characteristic	Pub: All (n=60)	DSIC: All (n=60)	Diff. (n=0)
Age (years)	10.9 (3.4-17.9)	10.9 (3.4-17.9)	0 (0)
Sex			
Girls	45 (75%)	45 (75%)	0 (0)
Boys	15 (25%)	15 (25%)	0 (0)
BMI (kg/m^2)	16.7 (13.8-30.6)	16.7 (13.8-30.6)	0 (0)
Race			
White	1 (2%)	1 (2%)	0 (0)
Black/African American	3 (5%)	3 (5%)	0 (0)
Asian	54 (90%)	54 (90%)	0 (0)
Mixed	2 (3%)	2 (3%)	0 (0)
Born in the United States/Canada			
Yes	10 (17%)	10 (17%)	0 (0)
No	50 (83%)	50 (83%)	0 (0)
HBV genotype			
A	1 (2%)	1 (2%)	0 (0)
B	32 (53%)	32 (53%)	0 (0)
C	22 (37%)	22 (37%)	0 (0)
D	3 (5%)	3 (5%)	0 (0)
E	2 (3%)	2 (3%)	0 (0)
HBV DNA (log_{10} IU/mL): Screening	8.3 (7.3-9.1)	8.3 (7.3-9.1)	0 (0)
HBV DNA (log_{10} IU/mL): Baseline	8.2 (7.5-9.1)	8.2 (7.5-9.1)	0 (0)
Quantitative HBeAg (log_{10} IU/mL)	3.2 (2.1-3.8)	3.2 (2.1-3.8)	0 (0)
Quantitative HBsAg (log_{10} IU/mL)	4.7 (3.2-5.4)	4.7 (3.2-5.4)	0 (0)
ALT (U/L): Screening	28 (10-57)	28 (10-57)	0 (0)
Boys	37 (23-57)	37 (23-57)	0 (0)
Girls	24 (10-40)	24 (10-40)	0 (0)
ALT $\leqslant$ new ULN: Screening	24 (40%)	24 (40%)	0 (0)
ALT (U/L): Baseline	29 (12-112)	29 (12-112)	0 (0)
Boys	39 (21-112)	39 (21-112)	0 (0)
Girls	26 (12-71)	26 (12-71)	0 (0)
ALT $\leqslant$ new ULN: Baseline	23 (38%)	23 (38%)	0 (0)
AST (U/L)	30 (14-92)	30 (14-92)	0 (0)
Total bilirubin (mg/dL)	0.4 (0.1-1.0)	0.4 (0.1-1.0)	0 (0)
Albumin (g/dL)	4.4 (3.5-5.1)	4.4 (3.5-5.1)	0 (0)
Platelets ($\times 10^3/mm^3$)	269 (134-492)	269 (134-492)	0 (0)
White blood cells ($\times 10^3/mm^3$)	5.8 (3.4-13.8)	5.8 (3.4-13.8)	0 (0)

Attachment A: SAS Code

```
libname dsic "X:\NIDDK\niddk-dr_studies2\HBRN\private_created_data\Pediatric IT\Redacted
Datasets";

/*************************/
/* DSIC for HBRN Peds IT */
/* Rosenthal et al. */
/*************************/

data long; set dsic.itp_long;
run;

*age;
proc means data=long n median min max;
var Age_visit;
where ittmpt = 2; *1 = screening, 2= baseline;
run;

*Sex;
proc freq data=long;
tables sex;
where ittmpt = 2;
run;

*BMI;
proc means data=long median min max;
var bmi;
where ittmpt = 2;
run;

*race;
proc freq data=long;
tables race;
where ittmpt = 2;
run;

*Born in the US;
proc freq data=long;
tables born_status;
where ittmpt = 2;
run;

*HBV genotype;
proc freq data=long;
tables cdcgen_cat;
where ittmpt = 2;
```

```
run;

*HBV DNA;
data long_1; set long;
hbvdna_log = log10(hbvdna_c);
run;

proc means data=long_1 median min max;
var hbvdna_log;
class ittmpt;
run;

*quant HBeAg and HBsAg;
data long_2; set long_1;
equant_log = log10(equant);
squant_log = log10(squant);
run;

proc means data=long_2 median min max;
var equant_log squant_log;
where ittmpt = 2;
run;

*alt screening;
proc sort data=long_2;
by sex;
run;

proc means data=long_2 median min max;
var alt;
by sex;
where ittmpt = 1;
run;

*alt LE new ULN screening;
data long_3; set long_2;
ULN_new = alt/alt_norm_new;
if alt <= ULN_new then alt_new_le = 1;
if alt > ULN_new then alt_new_le = 0;
run;

proc freq data=long_3;
tables alt_new_le;
where ittmpt = 1;
run;

*ALT baseline;
proc means data=long_3 median min max;
```

```
var alt;
by sex;
where ittmpt = 2;
run;

*alt LE new ULN baseline;
proc freq data=long_3;
tables alt_new_le;
where ittmpt = 2;
run;

*ast, total bilirubin, albumin, platelets, and wbc;
proc means data=long_3 median min max;
var ast tbili alb plat wbc;
where ittmpt = 2;
run;
```