Dataset Integrity Check for Characterization of ALT Flares During Treatment in the Immune Tolerant HBRN Trials (HBRN ALT Flares IT Trials)

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

An ancillary study from the Hepatitis B Research Network (HBRN) Immune Tolerant (IT) trials, the Characterization of ALT Flares During Treatment in the IT HBRN Trials, aimed to explore factors associated with the decline in viral antigens during the treatment of immune tolerant (IT) children and adults with combined peginterferon alfa-2a and entecavir.

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 Adult and Pediatric Immune Tolerant folders in the data packages. For this replication, variables were taken from datasets in the HBRN Adult and Pediatric Immune Tolerant trials (" and "itp_long.sas7bdat"), and merged with the ALT Flares IT Trials ancillary study dataset ("il28b_results.sas7bdat").

4 Statistical Methods

Analyses were performed to replicate results for the data in the publication by Perrillo 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 within expected variation to the published results.

6 Conclusions

The NIDDK Central Repository is confident that the HBRN ALT Flares IT Trials data files to be distributed are a true copy of the study data.

7 References

[1] Perrillo R, Lin HS, Schwarz KB, Rosenthal P, Lisker-Melman M, Chung RT, Prokunina-Olsson L, Cloherty G, Feld J. Changes in Serum Hepatitis B Surface and E Antigen, Interferon-inducible Protein 10, and Aminotransferase Levels during Combination Therapy of Immune-tolerant Chronic Hepatitis B. Hepatology, 76(3), 775-787, September 2022. doi: https://doi.org/10.1002/hep.32400

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

Table Variable	dataset.variable
Sex	it_long.sex
	itp_long.sex
Race	it_long.race
	itp_long.race
Age at baseline	it_long.age_visit
	itp_long.age_visit
HBV genotype	it_long.cdcgen_cat
	itp_long.cdcgen_cat
HBV DNA ($\log_{10}$ IU/ml)	it_long.hbvdna_c
	itp_long.hbvdna_c
HBsAg ($\log_{10}$ IU/ml)	it_long.squant
	itp_long.squant
HBeAg ($\log_{10}$ IU/ml)	it_long.equant
	itp_long.equant
ALT (U/L)	it_long.alt
	itp_long.alt
AST (U/L)	it_long.ast
	itp_long.ast
White blood cells ($\times 10^3$/mm^3)	it_long.wbc
	itp_long.wbc
Platelets ($\times 10^3$/mm^3)	it_long.plat
	itp_long.plat
IFNL4 (IL-28B) genotype	il28b_results.ifnl4_dg_tt

4

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

Characteristics	Pub: All (n=86)	DSIC: All (n=86)	Diff. (n=0)	Pub: Children (n=59)	DSIC: Children (n=59)	Diff. (n=0)	Pub: Adults (n=27)	DSIC: Adults (n=27)	Diff. (n=0)
Sex									
Male	29 (34%)	29 (34%)	0 (0)	15 (25%)	15 (25%)	0 (0)	14 (52%)	14 (52%)	0 (0)
Female	57 (66%)	57 (66%)	0 (0)	44 (75%)	44 (75%)	0 (0)	13 (48%)	13 (48%)	0 (0)
Race									
White	1 (1%)	1 (1%)	0 (0)	1 (2%)	1 (2%)	0 (0)	0 (0%)	0 (0%)	0 (0)
Black	4 (5%)	4 (5%)	0 (0)	3 (5%)	3 (5%)	0 (0)	1 (4%)	1 (4%)	0 (0)
Asian	79 (92%)	79 (92%)	0 (0)	53 (90%)	53 (90%)	0 (0)	26 (96%)	26 (96%)	0 (0)
Other/Mixed	2 (2%)	2 (2%)	0 (0)	2 (3%)	2 (3%)	0 (0)	0 (0%)	0 (0%)	0 (0)
Age at baseline	14.5 (8.3, 28.9) 3.4-61.2	14.5 (8.3, 28.9) 3.4-61.2	0 (0) 0	11.0 (7.3, 14.6) 3.4-17.9	11.0 (7.3, 14.6) 3.4-17.9	0 (0) 0	37.6 (30.1, 40.8) 22.2-61.2	37.6 (30.1, 40.8) 22.2-61.2	0 (0) 0
HBV genotype									
A	1 (1%)	1 (1%)	0 (0)	1 (2%)	1 (2%)	0 (0)	0 (0%)	0 (0%)	0 (0)
B	40 (47%)	40 (47%)	0 (0)	31 (53%)	31 (53%)	0 (0)	9 (33%)	9 (33%)	0 (0)
C	39 (45%)	39 (45%)	0 (0)	22 (37%)	22 (37%)	0 (0)	17 (63%)	17 (63%)	0 (0)
D	3 (3%)	3 (3%)	0 (0)	3 (5%)	3 (5%)	0 (0)	0 (0%)	0 (0%)	0 (0)
E	3 (3%)	3 (3%)	0 (0)	2 (3%)	2 (3%)	0 (0)	1 (4%)	1 (4%)	0 (0)
HBV DNA ($\log_{10}$ IU/ml)	8.2 (8.1, 8.4) 7.2-9.1	8.2 (8.1, 8.4) 7.2-9.1	0 (0) 0	8.2 (8.1, 8.5) 7.5-9.1	8.2 (8.1, 8.5) 7.5-9.1	0 (0) 0	8.3 (8.1, 8.4) 7.2-8.8	8.3 (8.1, 8.4) 7.2-8.8	0 (0) 0
HBsAg ($\log_{10}$ IU/ml)	4.7 (4.5, 4.8) 3.2-5.4	4.7 (4.5, 4.8) 3.2-5.4	0 (0) 0	4.7 (4.5, 4.8) 3.2-5.4	4.7 (4.5, 4.8) 3.2-5.4	0 (0) 0	4.7 (4.5, 4.8) 4.2-5.1	4.7 (4.5, 4.8) 4.2-5.1	0 (0) 0
HBeAg ($\log_{10}$ IU/ml)	3.3 (3.1, 3.4) 2.1-3.8	3.3 (3.1, 3.4) 2.1-3.8	0 (0) 0	3.2 (3.1, 3.4) 2.1-3.8	3.2 (3.1, 3.4) 2.1-3.8	0 (0) 0	3.3 (3.3, 3.4) 2.3-3.6	3.3 (3.3, 3.4) 2.3-3.6	0 (0) 0
ALT (U/L), male	33 (27, 41) 14-112	33 (27, 41) 14-112	0 (0) 0	39 (28, 43) 21-112	39 (28, 43) 21-112	0 (0) 0	29 (24, 33) 14-47	29 (24, 33) 14-47	0 (0) 0
ALT (U/L), female	23 (17, 32) 9-71	23 (17, 32) 9-71	0 (0) 0	27 (18, 35) 12-71	27 (18, 35) 12-71	0 (0) 0	15 (13, 19) 9-30	15 (13, 19) 9-30	0 (0) 0
AST (U/L)	27 (20, 33) 14-92	27 (20, 33) 14-92	0 (0) 0	30 (25, 38) 14-92	30 (25, 38) 14-92	0 (0) 0	20 (19, 24) 15-30	20 (19, 24) 15-30	0 (0) 0
White blood cells ($\times 10^3$/mm^3)	5.8 (4.7, 6.6) 3.0-13.8	5.8 (4.7, 6.6) 3.0-13.8	0 (0) 0	5.8 (5.2, 6.6) 3.4-13.8	5.8 (5.2, 6.6) 3.4-13.8	0 (0) 0	5.0 (4.1, 6.4) 3.0-9.1	5.0 (4.1, 6.4) 3.0-9.1	0 (0) 0
Platelets ($\times 10^3$/mm^3)	257 (211, 290) 134-492	257 (211, 290) 134-492	0 (0) 0	269 (212, 294) 134-492	269 (212, 294) 134-492	0 (0) 0	238 (209, 274) 137-325	238 (209, 274) 137-325	0 (0) 0
IFNL4 (IL-28B) genotype									
TT/TT	61 (87%)	60 (87%)	1 (0)	42 (88%)	41 (87%)	1 (1%)	19 (86%)	19 (86%)	0 (0)
dG/TT	9 (13%)	9 (13%)	0 (0)	6 (13%)	6 (13%)	0 (0)	3 (14%)	3 (14%)	0 (0)

Attachment A: SAS Code

```
libname adult_it "X:\NIDDK\niddk-dr_studies2\HBRN\private_created_data\Adult IT\Redacted Data";
libname ped_it "X:\NIDDK\niddk-dr_studies2\HBRN\private_created_data\Pediatric IT\Redacted
Datasets";
libname alt_it "X:\NIDDK\niddk-dr_studies2\HBRN\private_created_data\HBRN Ancillary Studies\HBRN
Ancillary Studies\ALT flares IT trials";

/**********************/
/* ALT Flares IT Trial */
/* DSIC Perrillo et al.*/
/**********************/

proc freq data=alt_it.il28b_results;
tables IFNL4_dG_TT;
run;

*identifying participants from the Ped and Adult IT trials;
data alt_one; set alt_it.cytokine_results;
keep orig_id id;
run;

data alt_two; set alt_it.ifn_antibody;
orig_id = id;
keep id orig_id;
run;

data alt_three; set alt_it.ifn_fold_change;
keep id orig_id;
run;

data alt_four; set alt_it.il28b_results;
orig_id = id;
keep id orig_id IFNL4_dG_TT;
run;

data combined; set
alt_one
alt_two
alt_three
alt_four;
run;

proc sort data=combined nodupkey;
by orig_id;
run;
```

```
*identifying the pediatric IT participants;
data ped; set ped_it.itp_long;
where ittmpt = 2;
run;

proc sort data=ped;
by orig_id;
run;

data alt_ped; merge
ped (in=a)
combined (in=b);
by orig_id;
if a=b;
run;

*identifying adult IT participants;
data adult; set adult_it.it_long;
where ittmpt = 2;
run;

proc sort data=adult;
by orig_id;
run;

data alt_adult; merge
adult (in=a)
combined (in=b);
by orig_id;
if a=b;
run;

/************************/
/* Combining Adult and Ped */
/* IT for combined alt_it dataset */
/************************/

data alt_ped_one; set alt_ped;
flag = 1;
keep orig_id sex race Age_visit cdcgen_cat hbvdna_c sQuant eQuant
alt ast wbc plat IFNL4_dG_TT keep flag;
run;

data alt_adult_one; set alt_adult;
flag = 2;
keep orig_id sex race Age_visit cdcgen_cat hbvdna_c sQuant eQuant
alt ast wbc plat IFNL4_dG_TT flag;
run;
```

```
data final_one; set
alt_ped_one
alt_adult_one;
run;

/************************/
/* Table 1 values */
/************************/

*sex;
proc freq data=final_one;
tables sex;
by flag;
run;

*race;
proc freq data=final_one;
tables race;
by flag;
run;

*age at baseline;
proc means data=final_one n median q1 q3 min max maxdec=2;
var Age_visit;
class flag;
run;

*HBV genotype;
proc freq data=final_one;
tables cdcgen_cat;
by flag;
run;

*HBV DNA;
data final_two; set final_one;
hbvdna = log10(hbvdna_c);
run;

proc means data=final_two n median q1 q3 min max maxdec=2;
var hbvdna;
class flag;
run;

*HBsAg and HBeAg;
data final_three; set final_one;
squant_log = log10(squant);
equant_log = log10(equant);
```

```
run;

proc means data=final_three n median q1 q3 min max maxdec=2;
var squant_log equant_log;
class flag;
run;

*ALT;
proc sort data=final_one;
by sex;
run;

proc means data=final_one n median q1 q3 min max maxdec=2;
var alt;
by sex;
class flag;
run;

*ast;
proc means data=final_one n median q1 q3 min max maxdec=1;
var ast;
class flag;
run;

*WBC;
proc means data=final_one n median q1 q3 min max maxdec=1;
var wbc;
class flag;
run;

*platelets;
proc means data=final_one n median q1 q3 min max maxdec=1;
var plat;
class flag;
run;

*IFNL4;
data ifnl; set alt_it.il28b_results;
keep id IFNL4_dG_TT;
where IFNL4_dG_TT ^= "";
run;

data ifnl_one; set ifnl;
orig_id = id;
run;

proc sort data=ifnl_one;
by orig_id;
```

```
run;

proc sort data=final_one;
by orig_id;
run;

data final_four; merge
final_one (in=a)
ifnl_one (in=b);
by orig_id;
if a=b;
run;

proc sort data=final_four dodupkey;
by orig_id;
run;

data final_five; set final_four;
if ifnl4_dg_TT = "Undet" then ifnl4_dg_tt = "";
run;

proc sort data=final_five;
by flag;
run;

proc freq data=final_five;
tables IFNL4_dG_TT;
by flag;
run;
```