

Dataset Integrity Check for  
Bone Density in Children With Chronic  
Liver Disease Correlates With Growth  
and Cholestasis – Techman et al

Prepared by Sabrina Chen

IMS Inc.

3901 Calverton Blvd, Suite 200 Calverton, MD 20705

April 16, 2019

## 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 1: Demographic Characteristics of A1AT Liver Disease Cohort, with Native Liver and Post-Transplant.....	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values.....	5
Attachment A: SAS Code .....	9

## 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 Childhood Liver Disease Research Network (ChiLDReN) is a National Institutes of Health–funded consortium of pediatric centers in North America focused on the study of rare pediatric liver diseases. The Longitudinal Study of Genetic Causes of Intrahepatic Cholestasis (LOGIC; NCT00571272) enrolled children with one of four diagnoses: ALGS, PFIC, A1AT, or bile acid synthetic disorder (BASD). As part of the LOGIC study protocol, participants aged >5 years underwent DXA scanning, the results of which are reported here. The objectives of this study were to address an important knowledge gap by investigating the prevalence of bone mineral deficits in this cohort of children with chronic liver disease and to determine factors associated with lower bone density, including growth parameters, laboratory values, and clinical events such as fracture and biliary diversion.

## 3 Archived Datasets

All the SAS data files, as provided by the Data Coordinating Center (DCC), are located in the LOGIC folder in the data package. For this replication, variables were taken from the “a1atbaseline.sas7bdat” dataset.

## 4 Statistical Methods

Analyses were performed to duplicate results for the data published by Techman et al [1] in the journal *J Pediatr Gastroenterol Nutr* in 2015. To verify the integrity of the dataset, descriptive statistics were computed.

## 5 Results

For Table 1 in the publication [1], **Demographic Characteristics of A1AT Liver Disease Cohort, with Native Liver and Post-Transplant**, 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.

## 6 Conclusions

The NIDDK repository is confident that the a1atbaseline data file to be distributed is a true copy of the study data.

## 7 References

[1] Techman J, et al. Baseline Analysis of a Young Alpha-1-AT Deficiency Liver Disease Cohort Reveals Frequent Portal Hypertension. *J Pediatr Gastroenterol Nutr*. 2015 July ; 61(1): 94–101.

**Table A:** Variables used to replicate Table 1: Demographic Characteristics of A1AT Liver Disease Cohort, with Native Liver and Post-Transplant

<b>Table Variable</b>	<b>dataset.variable</b>
Age, years	a1atbaseline.age
Age at symptoms first noticed, years	a1atbaseline.ihx2b01_age
Age at liver disease diagnosed, years	a1atbaseline.ihx2b02_age
Gender, N (%)	a1atbaseline.dmgab02
Ethnicity, N (%)	a1atbaseline.dmgab03
Race, N (%)	a1atbaseline.racecat
Genotype/phenotype, N (%)	a1atbaseline.sz
Genotype/phenotype, N (%)	a1atbaseline.zz
Cohort	a1atbaseline.pretx

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

<b>Character</b>	<b>All Manuscript</b>	<b>All DSIC</b>	<b>All Difference</b>	<b>Native Liver Manuscript</b>	<b>Native Liver DSIC</b>	<b>Native Liver Difference</b>	<b>Post-Transplant Manuscript</b>	<b>Post-Transplant DSIC</b>	<b>Post-Transplant Difference</b>
MEAN									
Age	8.2	8.2	0	6.9	6.9	0	12.9	12.9	0
Age at symptoms first noticed	1.3	1.3	0	1.4	1.4	0	0.9	0.9	0
Age at liver disease dx	1.7	1.7	0	1.8	1.8	0	1.3	1.3	0
SD									
Age	6.38	6.38	0	5.65	5.65	0	6.51	6.51	0
Age at symptoms first noticed	3.14	3.14	0	3.33	3.33	0	2.34	2.34	0
Age at liver disease dx	3.29	3.29	0	3.43	3.43	0	2.71	2.71	0
Median									
Age	6.8	6.8	0	5.5	5.5	0	13.7	13.7	0
Age at symptoms first noticed	0.1	0.1	0	0.1	0.1	0	0.2	0.2	0
Age at liver disease dx	0.3	0.3	0	0.3	0.3	0	0.2	0.2	0

Character	All Manuscript	All DSIC	All Difference	Native Liver Manuscript	Native Liver DSIC	Native Liver Difference	Post-Transplant Manuscript	Post-Transplant DSIC	Post-Transplant Difference
Q1									
Age	2.8	2.8	0	2.0	2	0	7.0	7	0
Age at symptoms first noticed	0	0	0	0	0	0	0.1	0.1	0
Age at liver disease dx	0.1	0.1	0	0.1	0.1	0	0.2	0.2	0
Q3									
Age	12.9	12.9	0	9.8	9.75	0.05	18.4	18.4	0
Age at symptoms first noticed	0.9	0.9	0	1.0	1	0	0.4	0.4	0
Age at liver disease dx	1.6	1.6	0	1.6	1.6	0	0.9	0.9	0
Min									
Age	0.0	0	0	0.0	0	0	0.4	0.4	0
Age at symptoms first noticed	0	0	0	0	0	0	0	0	0
Age at liver disease dx	0	0	0	0	0	0	0	0	0
Max									
Age	24.9	24.9	0	24.9	24.9	0	24.6	24.6	0

Character	All Manuscript	All DSIC	All Difference	Native Liver Manuscript	Native Liver DSIC	Native Liver Difference	Post-Transplant Manuscript	Post-Transplant DSIC	Post-Transplant Difference
Age at symptoms first noticed	20.2	20.2	0	20.2	20.2	0	12.6	12.6	0
Age at liver disease dx	20.2	20.2	0	20.2	20.2	0	12.9	12.9	0
N									
Age	269	269	0	208	208	0	61	61	0
Age at symptoms first noticed	245	245	0	190	190	0	55	55	0
Age at liver disease dx	255	255	0	199	199	0	56	56	0
Gender, N									
Male, N	164	164	0	124	124	0	40	40	0
Female, N	105	105	0	84	84	0	21	21	0
Ethnicity, N									
Hispanic	18	18	0	15	15	0	3	3	0
Non-Hispanic	247	247	0	189	189	0	58	58	0
Race, N									
Native American	1	1	0	1	1	0	0		0
Black	3	3	0	1	1	0	2	2	0
Caucasian	250	250	0	194	194	0	56	56	0
Multiracial	14	14	0	11	11	0	3	3	0

Character	All Manuscript	All DSIC	All Difference	Native Liver Manuscript	Native Liver DSIC	Native Liver Difference	Post-Transplant Manuscript	Post-Transplant DSIC	Post-Transplant Difference
Not Reported	1	1	0	1	1	0	0		0
Genotype/phenotype, N									
SZ	21	21	0	17	17	0	4	4	0
ZZ	246	246	0	190	190	0	56	56	0
Missing	2	2	0	1	1	0	1	1	0
Gender, %									
Male, %	61.0	61	0	59.6	59.6	0	65.6	65.6	0
Female, %	39.0	39	0	40.4	40.4	0	34.4	34.4	0
Ethnicity, %									
Hispanic	6.8	6.7	0.1	7.4	7.2	0.2	4.9	4.9	0
Non-Hispanic	93.2	91.8	1.4	92.6	90.9	1.7	95.1	95.1	0
Race, %									
Native American	0.4	0.4	0	0.5	0.5	0	0		0
Black	1.1	1.1	0	0.5	0.5	0	3.3	3.3	0
Caucasian	92.9	92.9	0	93.3	93.3	0	91.8	91.8	0
Multiracial	5.2	5.2	0	5.3	5.3	0	4.9	4.9	0
Not Reported	0.4	0.4	0	0.5	0.5	0	0		0
Genotype/phenotype, %									
SZ	7.8	7.8	0	8.2	8.2	0	6.6	6.6	0
ZZ	91.5	91.4	0.1	91.4	91.3	0.1	91.8	91.8	0
Missing	0.7	0.7	0	0.5	0.5	0	1.6	1.6	0

## Attachment A: SAS Code

```
options nocenter validvarname=upcase;

title '/prj/niddk/ims_analysis/LOGIC/prog_initial_analysis/dsic.LOGIC_Techman.y2019m04d15.sas';
run;

*****;
* INPUT ;
*****;
libname orig '/prj/niddk/ims_analysis/LOGIC/private_orig_data/Teckman_LOGIC_PROBE_A1AT/';

*****;
* MACROS ;
*****;
%macro readin(ds);
  data &ds;
    set sasfile.&ds;
  run;

  proc contents data=&ds;
    title3 "&ds";
  run;
%mend;

* produce n and %;
%macro npercent(rownum, var, varf, subset, subsetname);
  proc freq data=analy noprint;
    where &subset = 1;
    tables &var/list missing out=tbl1&subsetname;
  run;

  data tbl1&subsetname;
    length covar covarf $100;
    set tbl1&subsetname;
    covar = "&var";
    covarf = put(&var,&varf.);
    rownum = &rownum;
  run;

  data prnt&subsetname;
    set prnt&subsetname tbl1&subsetname;
  run;
%mend;
```

```

%macro univ(rownum, var, subset, subsetname);

proc univariate data=analy outtable= univ&subsetname noprint;
  where &subset=1;
  var &var
  ;
run;

data univ&subsetname;
  length covarf $100 _var_ $25;
  set univ&subsetname;
  covarf = "&subset";
  rownum = &rownum;
run;

data prntuniv&subsetname;
  set prntuniv&subsetname univ&subsetname;
run;

%mend;

```

```

*****;
* FORMATS      ;
*****;
proc format;
  value novalue
    . = "No Value"
  other = " Value"
  ;

  value genderf
    1=' Male'
    2='Female'
  ;

  value ethnicf
    1 = 'Hispanic'
    2 = 'Non-hispanic'
    66 = 'Not reported'
  ;

  value racef
    1 = ' Native American'
    2 = 'Asian'
    3 = 'Black or African-American'
    4 = 'Native Hawaiian or PI'
    5 = 'Caucasian'
    6 = 'Unknown'

```

```

7 = 'Refused'
8 = 'Multiracial'
;

value genof
1 = " SZ"
2 = " ZZ"
99 = "Missing"
;
run;

data baseline;
  set orig.alatbaseline;
run;

proc contents data=baseline;
run;

proc freq data=baseline;
  table pretx/missing;
run;

data analy;
  set baseline;
  * create subset flag for each row to use in macro call;
  all = 1;

  if pretx = 1 then in_native=1;
  else if pretx = 0 then in_post = 1;

  if sz=1 and zz=0 then geno_sz_zz = 1;
  else if sz=0 and zz=1 then geno_sz_zz = 2;
  else if sz=0 and zz=0 then geno_sz_zz = 99;

  if racecat='A'      then racecatn = 1;
  else if racecat='B' then racecatn = 2;
  else if racecat='C' then racecatn = 3;
  else if racecat='D' then racecatn = 4;
  else if racecat='E' then racecatn = 5;
  else if racecat='F' then racecatn = 6;
  else if racecat='G' then racecatn = 7;
  else if racecat='H' then racecatn = 8;

  if dmgab02 = '1' then dmgab02n = 1;
  else if dmgab02 = '2' then dmgab02n = 2;

  if dmgab03 = '1' then dmgab03n = 1;
  else if dmgab03 = '2' then dmgab03n = 2;
  else if dmgab03 = '66' then dmgab03n = 66;

```

```

run;

proc freq data=analy;
  tables pretx*in_native*in_post/list missing;
  tables geno_sz_zz*sz*zz/list missing;
  tables racecat*racecatn/list missing;
  tables dmgab02*dmgab02n
         dmgab03*dmgab03n/list missing;
run;

* med, min and max;
data prntunivall;
  set _null_;
run;

%univ(1, age          , all      , all);
%univ(2, ihx2b01_age , all      , all);
%univ(3, ihx2b02_age , all      , all);

proc print data= prntunivall noobs;
  var rownum _var_ covarf _nobs_ _mean_ _std_ _median_ _q1_ _q3_ _min_ _max_ ;
run;

data prntunivnative;
  set _null_;
run;

%univ(1, age          , in_native  , native);
%univ(2, ihx2b01_age , in_native  , native);
%univ(3, ihx2b02_age , in_native  , native);

proc print data= prntunivnative noobs;
  var rownum _var_ covarf _nobs_ _mean_ _std_ _median_ _q1_ _q3_ _min_ _max_ ;
run;

data prntunivposttransplant;
  set _null_;
run;

%univ(1, age          , in_post      , posttransplant);
%univ(2, ihx2b01_age , in_post      , posttransplant);
%univ(3, ihx2b02_age , in_post      , posttransplant);

proc print data= prntunivposttransplant noobs;
  var rownum _var_ covarf _nobs_ _mean_ _std_ _median_ _q1_ _q3_ _min_ _max_ ;
run;

* combine rows;

```

```

proc sort data=prntunivnative (rename=( _nobs_ = nav_nobs_
      _median_ = nav_median_
      _q1_ = nav_q1_
      _q3_ = nav_q3_
      _min_ = nav_min_
      _max_ = nav_max_
      _mean_ = nav_mean_
      _std_ = nav_std_ ))
;
  by rownum covarf;
run;

proc sort data=prntunivposttransplant (rename=( _nobs_ = post_nobs_
      _median_ = post_median_
      _q1_ = post_q1_
      _q3_ = post_q3_
      _min_ = post_min_
      _max_ = post_max_
      _mean_ = post_mean_
      _std_ = post_std_ ))
;
  by rownum covarf;
run;

data alluniv;
  merge prntunivall      (in=in1 keep = rownum _var_ covarf _nobs_ _mean_ _std_ _median_ _q1_ _q3_ _min_ _max_ )
        prntunivnative  (in=in2 keep = rownum _var_ covarf nav_nobs_ nav_median_ nav_q1_ nav_q3_ nav_min_ nav_max_ nav_mean_
nav_std_ )
        prntunivposttransplant (in=in3 keep = rownum _var_ covarf post_nobs_ post_median_ post_q1_ post_q3_ post_min_ post_max_
post_mean_ post_std_ );
  by rownum;
  if in1 or in2 or in3;
run;

* n and percent;
data prntall;
  set _null_;
run;

%npercent(4, dmgab02n , genderf , all , all );
%npercent(5, dmgab03n , ethnicf , all , all );
%npercent(6, racecatn , racef , all , all );
%npercent(7, geno_sz_zz , genof , all , all );

proc print data=prntall;
  var rownum covarf COUNT PERCENT;
title3 "All Cohort";
run;

```

```

data prntnative;
  set _null_;
run;

%npercent(4, dmgab02n , genderf , in_native , native );
%npercent(5, dmgab03n , ethnicf , in_native , native );
%npercent(6, racecatn , racef , in_native , native );
%npercent(7, geno_sz_zz , genof , in_native , native );

proc print data=prntnative;
  var rownum covarf COUNT PERCENT;
  title3 "Native Liver Cohort";
run;

data prntposttransplant;
  set _null_;
run;

%npercent(4, dmgab02n , genderf , in_post , posttransplant);
%npercent(5, dmgab03n , ethnicf , in_post , posttransplant);
%npercent(6, racecatn , racef , in_post , posttransplant);
%npercent(7, geno_sz_zz , genof , in_post , posttransplant);

proc print data=prntposttransplant;
  var rownum covarf COUNT PERCENT;
  title3 "Post-Transplant";
run;

* combine rows;
proc sort data=prntall;
  by rownum covarf;
run;

proc sort data=prntnative (rename=(count = natv_count
                                percent = natv_percent));
  by rownum covarf;
run;

proc sort data=prntposttransplant (rename=(count = post_count
                                           percent = post_percent));
  by rownum covarf;
run;

data allnprcnt;
  merge prntall (in=in1 keep=rownum covarf count percent)
        prntnative (in=in2 keep=rownum covarf natv_count natv_percent)
        prntposttransplant (in=in3 keep=rownum covarf post_count post_percent);
  by rownum covarf;
  if in1 or in2 or in3;

```

```
run;
```

```
data table1;
```

```
  set alluniv allnprcnt;  
  _mean_      = round(_mean_      ,.1);  
  nav_mean_   = round(nav_mean_   ,.1);  
  post_mean_  = round(post_mean_  ,.1);  
  _std_       = round(_std_       ,.01);  
  nav_std_    = round(nav_std_    ,.01);  
  post_std_   = round(post_std_   ,.01);  
  percent     =round(percent, .1);  
  natv_percent =round(natv_percent, .1);  
  post_percent =round(post_percent, .1);
```

```
  label _nobs_      = 'All N          '  
        nav_nobs_   = 'Native Liver N      '  
        post_nobs_  = 'Post-Transplant N   '  
        _mean_      = 'All Mean          '  
        nav_mean_   = 'Native Liver Mean   '  
        post_mean_  = 'Post-Transplant Mean '  
        _std_       = 'All              SD  '  
        nav_std_    = 'Native Liver      SD  '  
        post_std_   = 'Post-Transplant SD  '  
        _median_    = 'All              Median '  
        nav_median_ = 'Native Liver      Median '  
        post_median_ = 'Post-Transplant Median '  
        _q1_        = 'All              Q1   '  
        nav_q1_     = 'Native Liver      Q1   '  
        post_q1_    = 'Post-Transplant Q1   '  
        _q3_        = 'All              Q3   '  
        nav_q3_     = 'Native Liver      Q3   '  
        post_q3_    = 'Post-Transplant Q3   '  
        _min_       = 'All              Min  '  
        nav_min_    = 'Native Liver      Min  '  
        post_min_   = 'Post-Transplant Min  '  
        _max_       = 'All              Max  '  
        nav_max_    = 'Native Liver      Max  '  
        post_max_   = 'Post-Transplant Max  '  
        count       = 'All              N    '  
        natv_count  = 'Native Liver      N    '  
        post_count  = 'Post-Transplant N    '  
        percent     = 'All              Percent '  
        natv_percent = 'Native Liver      Percent '  
        post_percent = 'Post-Transplant Percent '  
  ;
```

```
run;
```

```
proc print data=table1 label;  
  var rownum _var_ covarf  
      _nobs_
```

```
nav_nobs_  
post_nobs_  
_mean_  
nav_mean_  
post_mean_  
_std_  
nav_std_  
post_std_  
_median_  
nav_median_  
post_median_  
_q1_  
nav_q1_  
post_q1_  
_q3_  
nav_q3_  
post_q3_  
_min_  
nav_min_  
post_min_  
_max_  
nav_max_  
post_max_  
  
count  
natv_count  
post_count  
percent  
natv_percent  
post_percent;  
run;  
  
endsas;  
alatbaseline.age  
alatbaseline.ihx2b01_age  
alatbaseline.ihx2b02_age  
alatbaseline.dmgab02  
alatbaseline.dmgab03  
alatbaseline.racecat  
alatbaseline.geno_sz_zz
```