Dataset Integrity Check for Adolescent Bariatrics: Assessing Health Benefits and Risks (Teen-LABS) Final Submission

Prepared by NIDDK-CR
February 21, 2024

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

Adolescent Bariatrics: Assessing Health Benefits and Risks (Teen-LABS), also known as Teen-Longitudinal Assessment of Bariatric Surgery, proposed that bariatric surgery was more beneficial to extremely obese participants when done during the adolescent years instead of adulthood. By using duration of obesity as the moderating variable, the Teen-LABS study estimated the risks and benefits of bariatric surgery among adolescent participants compared with adult participants. Over 200 adolescent bariatric patients were recruited from four centers and underwent gastric bypass surgery between 2007 and 2012. Post-surgery data and biospecimens were obtained at pre-determined points during a 24-month period. The assessments of the Teen-LABS participants were compared with similar data from the adult participants of the Longitudinal Assessment of Bariatric Surgery (LABS) study.

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 Teen-LABS Final folder in the data package. For this replication, variables were taken from the “ef.sas7bdat”, “gs.sas7bdat”, “ryb.sas7bdat”, “swh.sas7bdat”, “cdi.sas7bdat”, “anth.sas7bdat”, “cab.sas7bdat”, and “iwqol.sas7bdat” datasets from the Teen-LABS Final submission.

4 Statistical Methods

Analyses were performed to replicate results for the data in the publication by Ogel 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 by Age at Surgery, 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 Teen-LABS Final data files to be distributed are a true copy of the study data.

7 References

[1] Ogle SB, Dewberry LC, Jenkins TM, Inge TH, Kelsey M, Bruzoni M, Pratt JSA. Outcomes of Bariatric Surgery in Older Versus Younger Adolescents. Pediatrics, 147(3), e2020024182, March 2021. DOI: https://doi.org/10.1542/peds.2020-024182

Table A: Variables used to replicate Table 1 – Baseline Characteristics by Age at Surgery

Table Variable	**dataset.variable**
Age at surgery	ef.dob gs.surgdat ryb.surgdat
Female sex	ef.sex
White	ef.racew
Non-Hispanic	ef.ethn
Surgical procedure	ryb gs
Household income	swh.hincome cdi.pcgincome
Caregiver education attained	cdi.pcgeduc
Weight	anth.wgt1
BMI	anth.wgt1 anth.hgt1
Systolic blood pressure	anth.sbp1
Diastolic blood pressure	anth.dbp1
Type 2 diabetes	cab.dmconfirm
Hypertension	cab.hyptn
Dyslipidemia	cab.dyslipid
IWQOL total score	iwqol

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

Characteristic	Publication: Ages 13-15 years (n=66)	DSIC: 13-15 years (n=62)	Diff. (n=4)	Publication: Ages 16-19 years (n=162)	DSIC: Ages 16-19 years (n=158)	Diff. (n=4)
Age at surgery, mean (SD), years	15.1 (0.81)	15.1 (0.80)	0 (0.01)	17.7 (1.04)	17.7 (1.05)	0 (0.01)
Female sex, % (n)	72.7 (48)	72.6 (45)	0.1 (3)	75.9 (123)	75.9 (120)	0 (3)
White, % (n)	72.7 (48)	79.0 (49)	6.3 (1)	71.6 (116)	76.6 (121)	5 (5)
Non-Hispanic, % (n)	97.0 (64)	98.4 (61)	1.4 (3)	91.4 (148)	91.1 (144)	0.3 (4)
Surgical procedure, % (n)						
RYGB	63.6 (42)	64.5 (40)	0.9 (2)	73.5 (119)	74.0 (117)	0.5 (2)
VSG	36.4 (24)	35.5 (22)	0.9 (2)	26.5 (43)	25.9 (41)	0.6 (2)
Household income, % (n)						
< $25,000	34.4 (22)	33.3 (20)	1.1 (2)	39.6 (61)	38.3 (59)	1.3 (2)
$25,000 to $74,999	34.4 (22)	36.7 (22)	2.3 (0)	39.0 (60)	39.0 (60)	0 (0)
$75,000+	31.3 (20)	30.0 (18)	1.3 (2)	21.4 (33)	22.7 (35)	1.3 (2)
Caregiver education attained, % (n)						
Less than high school	10.8 (7)	9.7 (6)	1.1 (1)	10.3 (16)	12.7 (20)	2.4 (4)
High school graduate	30.8 (20)	30.6 (19)	0.2 (1)	30.8 (48)	28.5 (45)	2.3 (3)
Some college	41.5 (27)	41.9 (26)	0.4 (1)	39.7 (62)	39.9 (63)	0.2 (1)
College graduate or higher	16.9 (11)	17.7 (11)	0.8 (0)	19.2 (30)	19.0 (30)	0.2 (0)
Weight, mean (SD), kg	148.6 (31.88)	149.6 (32.42)	1.0 (0.54)	148.9 (31.05)	148.2 (32.79)	0.7 (1.74)
BMI, mean (SD)	53.1 (10.75)	53.4 (10.94)	0.3 (0.19)	52.4 (8.79)	52.5 (8.64)	0.1 (0.15)
Systolic blood pressure, mean (SD), mm Hg	121.5 (12.53)	120.7 (21.53)	0.8 (9.0)	126.8 (13.31)	125.6 (22.14)	1.2 (8.83)
Diastolic blood pressure, mean (SD), mm Hg	71.2 (8.96)	70.3 (13.89)	0.9 (4.93)	75.4 (10.27)	74.3 (15.19)	1.1 (4.92)
Type 2 diabetes, % (n)	10.6 (7)	11.3 (7)	0.7 (0)	13.6 (22)	15.2 (24)	1.6 (2)
Hypertension, % (n)	27.3 (18)	25.8 (16)	1.5 (2)	37.1 (59)	29.3 (46)	7.8 (13)
Dyslipidemia, % (n)	73.4 (47)	41.0 (25)	32.4 (22)	77.0 (124)	48.7 (77)	28.3 (47)
IWQOL total score, mean (SD)	64.1 (17.75)	96.5 (19.13)	32.4 (1.38)	61.7 (18.16)	93.3 (19.51)	31.6 (1.35)

Attachment A: SAS Code

```
libname final "X:\NIDDK\niddk-dr_studies6\Teen_Labs\private_created_data\Year_8_AND_Final for DEID\TeenLABS_FINAL\DB\Upto15yr";
libname same "X:\NIDDK\niddk-dr_studies6\Teen_Labs\private_created_data\Year_8_AND_Final for DEID\TeenLABS_FINAL\DB\SameVisit";

/*****************************/
/* Teen LABS Final Submission DSIC*/
/*****************************/

*Identifying variables to use for age calculation;
proc freq data=final.ef;
tables dob;
run;

proc freq data=same.gs;
tables surgdat;
run;

proc freq data=same.ryb;
tables surgdat;
run;

*Creating separate dataset to identify age groups used in publication;
data ryb; set same.ryb;
surgtype = 1;
keep id_new surgdat keep surgtype;
run;

data gs; set same.gs;
surgtype = 2;
keep id_new surgdat surgtype;
run;

data ef; set final.ef;
keep id_new dob;
run;

*Merging datasets;
proc sort data=ryb;
by id_new;
run;

proc sort data=gs;
by id_new;
run;
```

```
proc sort data=ef;
by id_new;
run;

data age_grp; merge
ryb (in=a)
gs (in=b)
ef (in=c);
by id_new;
if a=c or b=c;
run;

proc freq data=age_grp;
tables surgdat;
run;

*Reducing participants to those getting surgery between 2007 and 2012;
data age_grp1; set age_grp;
if '01Jan2007'd <= surgdat <= '31Dec2012'd;
run;

*Creating an age variable;
data age_grp_2; set age_grp1;
age = (surgdat - dob)/365.25;
run;

proc freq data=age_grp_2;
tables age;
run;

*Creating the flag variable for the groups;
data age_grp3; set age_grp_2;
group = 0;
if age > 13 AND age < 16 then group=1;
if age >=16 AND age < 20 then group=2;
run;

proc freq data=age_grp3;
tables group;
run;

*Identifying those that made it to 5 years follow-up;
proc freq data=final.foa;
tables visit;
run;

data foa; set final.foa;
```

```
if visit <= 72;
keep id_new;
run;

proc sort data=foa nodup;
by id_new;
run;

*Merging the foa ids with the age group data;
data age_grp4; merge
foa (in=a)
age_grp3 (in=b);
by id_new;
if a=b;
run;

proc freq data=age_grp4;
tables group;
run;
*

/*******************************/
*Replicating table 1 from pub;
/*******************************/

*Mean age at surgery;
proc means data=age_grp4 n mean std;
var age;
class group;
run;

*Sex;
data ef1; set final.ef;
run;

proc sort data=ef1;
by id_new;
run;

data one; merge
ef1 (in=a)
age_grp4 (in=b);
by id_new;
if a=b;
run;

proc freq data=one;
tables sex*group/norow nopercent;
```

```
run;

*Race (white);
proc freq data=one;
tables racew*group/norow nopercent missing;
run;

*Ethnicity (non-Hispanic);
proc freq data=one;
tables ethn*group/norow nopercent missing;
run;

*Surgery type;
proc sort data=ryb nodupkey;
by id_new;
run;

proc sort data=gs nodupkey;
by id_new;
run;

data two; merge
ryb (in=a)
gs (in=b)
one (in=c);
by id_new;
if c=1;
run;

proc freq data=two;
tables surgtype*group/norow nopercent missing;
run;

*Household income;
data swh; set final.swh;
if visit=1;
keep id_new hincome;
run;

data cdi; set final.cdi;
if visit=1;
hincome = pcghincome;
keep id_new hincome;
run;

proc sort data=swh;
by id_new;
run;
```

```
proc sort data=cdi;
by id_new;
run;

data income; merge
swh (in=a)
one (in=b)
cdi (in=c);
by id_new;
if b=1;
run;

proc freq data=income;
tables hincome;
run;

*Need to adjust the values for income;
/*-3=Don't know;
1.1=Less than $5,000;
1.2=$5,000 - $14,999;
1.3=$15,000 - $24,999;
2=$25,000 - $49,999;
3=$50,000 - $74,999;
4=$75,000 - $99,999;
5=$100,000 - $199,999;
6=$200,000 or more*/

data income1; set income;
houseinc = .;
if hincome <=1.3 AND hincome >= 0 then houseinc = 1;
if hincome = 2 OR hincome = 3 then houseinc = 2;
if hincome >=4 then houseinc = 3;
run;

proc freq data=income1;
tables houseinc*group/norow nopercent;
run;

*Education attained by caregiver;
proc freq data=final.cdi;
tables pcgeduc;
where visit = 1;
run;

data edu; set final.cdi;
if visit = 1;
keep id_new pcgeduc;
```

```
run;

proc sort data=edu;
by id_new;
run;

data educ; merge
one (in=a)
edu (in=b);
by id_new;
if a=1;
run;

/*"1=Less than high school;
2=Some high school (grades 9-12, no diploma or GED);
3=Some home-schooling (grades 9-12, no diploma or GED);
4=General Equivalency Degree (GED);
5=Graduated from high school;
6=1 to 2 years of college, no degree yet;
7=3 or more years of college, no degree yet;
8=Graduated from a 2-year college, business or vocational school, or got an Associate's degree;
9=Graduated from a college university and obtained a Bachelor's degree (BS, BA);
10=Some graduate school courses;
11=Master's degree;
12=Professional degree: Ph.D., Psy.D., Ed.D. M.D., DDS, LLB, LLD, JD etc."
*/

data educ1; set educ;
educ = .;
if pcgeduc < 4 then educ = 1;
if pcgeduc = 4 OR pcgeduc = 5 then educ = 2;
if pcgeduc = 6 OR pcgeduc = 7 OR pcgeduc = 8 then educ = 3;
if pcgeduc > 8 then educ = 4;
run;

proc freq data=educ1;
tables educ*group/norow nopercent;
run;

*Weight;
data anth; set final.anth;
if visit = 1;
keep id_new wgt1 wgt2 wgt3 hgt1 hgt2 hgt3 sbp1 sbp2 sbp3 dbp1
dbp2 dbp3;
run;

proc sort data=anth;
by id_new;
```

```
run;

data wgt; merge
one (in=a)
anth (in=b);
by id_new;
if a=1;
run;

proc means data=wgt n mean std;
var wgt wgt1 wgt2 wgt3;
class group;
run;

*BMI;
data bmi; set wgt;
bmi = .;
hgt_bmi = hgt1/100;
bmi = (wgt1/(hgt_bmi*hgt_bmi));
run;

proc means data=bmi n mean std;
var bmi;
class group;
run;

*SBP and DBP;
proc means data=bmi n mean std;
var sbp1 dbp1;
class group;
run;

*Type 2 diabetes;
proc freq data=final.cab;
tables dmconfirm;
run;

data dm; set final.cab;
keep id_new dmconfirm;
if visit = 1;
run;

proc sort data=dm;
by id_new;
run;

data dm1; merge
dm (in=a)
```

```
one (in=b);
by id_new;
if b=1;
run;

proc freq data=dm1;
tables dmconfirm*group/norow nopercent;
run;

*Hypertension and dyslipidemia;
data hptn; set final.cab;
keep id_new hyptn dyslipid;
run;

proc sort data=hptn;
by id_new;
run;

data hptn1; merge
one (in=a)
hptn (in=b);
by id_new;
if a=1;
run;

data hptn2; set hptn1;
htn = .;
if hyptn = 0 then htn = 0;
if hyptn > 0 then htn = 1;
dysl=.;
if dyslipid = 0 then dysl = 0;
if dyslipid > 0 then dysl = 1;
run;

proc freq data=hptn2;
tables (htn dysl)*group/norow nopercent;
run;

*IWQOL total score;
proc freq data=final.iwqol;
run;

data iwqol; set final.iwqol;
if visit = 1;
run;

proc sort data=iwqol;
by id_new;
```

```
run;

data iwqol1; merge
one (in=a)
iwqol (in=b);
by id_new;
if a=1;
run;

data iwqol2; set iwqol1;
totscore = IWQOL1 + IWQOL2 + IWQOL3 + IWQOL4 + IWQOL5 + IWQOL6
+ IWQOL7 + IWQOL8 + IWQOL9 + IWQOL10 + IWQOL11 + IWQOL12 + IWQOL13
+ IWQOL14 + IWQOL15 + IWQOL16 + IWQOL17 + IWQOL18 + IWQOL19
+ IWQOL20 + IWQOL21 + IWQOL22 + IWQOL23 + IWQOL24 + IWQOL25
+ IWQOL26 + IWQOL27;
run;

proc means data=iwqol2 n mean std;
var totscore;
class group;
run;
```