Dataset Integrity Check for Adolescent Bariatrics: Assessing Health Benefits and Risks (Teen-LABS) Year 8 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 Year 8 folder in the data package. For this replication, variables were taken from the “ef.sas7bdat”, “foa.sas7bdat”, “anth.sas7bdat”, “lv.sas7bdat”, and “central_lab.sas7bdat” datasets from Teen-LABS Year 8.

4 Statistical Methods

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

5 Results

For Table 1 in the publication [1], Adjusted Clinical Variables at Baseline and 5-Year Follow-up, According to Study Group, 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 Year 8 data files to be distributed are a true copy of the study data.

7 References

[1] Inge TH, Courcoulas AP, Jenkins TM, Michalsky MP, Brandt ML, Xanthakos SA, Dixon JB, Harmon CM, Chen MK, Xie C, Evans ME, Helmrath MA. Five-Year Outcomes of Gastric Bypass in Adolescents as Compared with Adults. The New England Journal of Medicine, 380(22), 2136-2145, May 2019. DOI: https://doi.org/10.1056/NEJMoa1813909

Table A: Variables used to replicate Table 1 – Adjusted Clinical Variables at Baseline and 5-Year Follow-up, According to Study Group

Table Variable	**dataset.variable**
Age	ef.dob (assumes 15th of birth month) ef.doctldat foa.foadat
Female sex	ef.sex
BMI	anth.wgt1 anth.wgt2 anth.wgt3 anth.hgt1 anth.hgt2 anth.hgt3
Weight	anth.wgt1 anth.wgt2 anth.wgt3
Systolic blood pressure	anth.sbp1 anth.sbp2 anth.sbp3
Diastolic blood pressure	anth.dbp1 anth.dbp2 anth.dbp3
Glycated hemoglobin	lv.hba1c cenral_lab.lab_hba1c
HDL cholesterol	lv.hdl central_lab.lab_hdl
Non-LDL cholesterol	lv.tc lv.hdl central_lab.chol central_lab.hdl
LDL cholesterol	lv.ldl central_lab.lab_ldl
Triglyceride	lv.trig central_lab.lab_TG_NET

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

Characteristic	Adolescents					
	Publication: Baseline (n=161)	**DSIC: Baseline (n=162)**	**Diff. (n=1)**	**Publication: At 5 Years (n=141)**	**DSIC: At 5 Years (n=138)**	**Diff. (n=3)**
Age (years)						
Range	13 to 19	13 to 20	0-1	18 to 25	18 to 25	0-0
Mean (95% CI)	17.0 (16.8 to 17.2)	17.0 (16.7 to 17.2)	0 (0.1-0)	22.1 (21.8 to 22.4)	22.1 (21.8 to 22.3)	0 (0-0.1)
Female sex (%)	78	78	0	79	80	1
BMI						
Mean (95% CI)	50 (47 to 52)	53 (52 to 55)	3 (5-3)	37 (35 to 40)	40 (39 to 43)	3 (4-3)
Weight (kg)						
Mean (95% CI)	147 (139 to 154)	149 (144 to 154)	2 (5-0)	111 (103 to 118)	115 (109 to 121)	4 (6-3)
Systolic blood pressure (mm Hg)						
Mean (95% CI)	127 (123 to 131)	125 (122 to 129)	2 (1-2)	121 (117 to 125)	121 (119 to 124)	0 (2-1)
Diastolic blood pressure (mm Hg)						
Mean (95% CI)	75 (72 to 78)	75 (72 to 77)	0 (0-1)	73 (70 to 76)	74 (72 to 76)	1 (2-0)
Glycated hemoglobin (%)						
Mean (95% CI)	5.2 (4.9 to 5.5)	5.9 (5.6 to 6.4)	0.7 (0.7-0.9)	5.1 (4.8 to 5.3)	5.3 (5.1 to 5.5)	0.2 (0.3-0.2)
HDL cholesterol (mg/dL)						
Mean (95% CI)	39 (35 to 43)	40 (38 to 41)	1 (3-2)	56 (51 to 60)	55 (52 to 57)	1 (1-3)
Non-LDL cholesterol (mg/dL)						
Mean (95% CI)	117 (107 to 127)	130 (123 to 136)	13 (16-9)	97 (88 to 107)	98 (93 to 103)	1 (5-4)
LDL cholesterol (mg/dL)						
Mean (95% CI)	89 (81 to 98)	104 (98 to 110)	15 (17-12)	80 (72 to 89)	82 (78 to 86)	2 (6-3)
Triglyceride (mg/dL)						
Mean (95% CI)	121 (105 to 139)	133 (118 to 148)	12 (13-9)	73 (63 to 84)	80 (72 to 88)	7 (9-4)

Attachment A: SAS Code

```
libname yr8 "X:\NIDDK\niddk-dr_studies6\Teen_Labs\private_created_data\Year_8_AND_Final for DEID\Year 8";

/*****************************/
/* Teen-LABS Year 8 DSIC */
/*****************************/

*Demographics data from enrollment form;
data enrol; set yr8.ef;
run;

proc contents data=yr8.ryb;
run;

*RCAB baseline form;
data rcab; set yr8.rcab;
run;

*Anthropometrics;
data anth1; set yr8.anth;
where visit = 1;
run;

*Only including participants who underwent gastric bypass;
data ryb; set yr8.ryb;
keep id_new ;
run;

*Merging to include only ryb participants;
proc sort data=ryb;
by id_new;
run;

proc sort data=enrol;
by id_new;
run;

data one; merge
enrol (in=a)
ryb (in=b);
by id_new;
if a=b;
run;

*Age;
```

```
data two; set one;
age = (doctldat - dob)/365.25;
run;

proc freq data=two;
tables age;
run;

proc means data=two n min max mean clm;
var age;
run;

*Sex;
proc freq data=two;
tables sex;
run;

*BMI;
data rcab1; set rcab;
keep id_new prebmi;
run;

proc sort data=rcab1;
by id_new;
run;

proc sort data=anth1;
by id_new;
run;

data three; merge
rcab1 (in=a)
two (in=b)
anth1 (in=c) ;
by id_new;
if a=b=c;
run;

*Calculating BMI using anth data;
proc freq data=anth1;
tables wgt1 wgt2 wgt3 hgt1 hgt2 hgt3;
run;

data four; set three;
avg_wgt = (wgt1+wgt2+wgt3)/3;
avg_hgt = ((hgt1+hgt2+hgt3)/3)/100;

calc_bmi= avg_wgt/(avg_hgt*avg_hgt);
```

```
run;

proc freq data=four;
tables avg_hgt avg_wgt calc_bmi;
run;

*The two variables within the data for BMI are slightly off from the pub, calculating BMI manually from
the ANTH dataset yields unrealistic numbers, the weight values that are recorded seem high for
adolescents (e.g., 254kg);
proc means data=four n mean clm;
var efbmi prebmi calc_bmi;
run;

*Weight;
proc means data=four n mean clm;
var avg_wgt wgt wgt1 wgt2 wgt3;
run;

*SBP;
proc means data=four n mean clm;
var sbp1 sbp2 sbp3;
run;

*DBP;
proc means data=four n mean clm;
var dbp1 dbp2 dbp3;
run;

*Glycated hemoglobin (A1c);
data lv; set yr8.lv;
where Visit = 1;
run;

proc sort data=lv;
by id_new;
run;

data five; merge
four (in=a)
lv (in=b);
by id_new;
if a=b;
run;

proc means data=five n mean clm;
var hba1c;
where hba1c > 0;
run;
```

```
*HDL cholesterol;
proc means data=five n mean clm;
var hdl;
where hdl > 0;
run;

*Non-LDL cholesterol;
data six; set five;
nonldl = tc - hdl;
run;

proc means data=six n mean clm;
var nonldl;
where tc > 0;
run;

*LDL;
proc means data=six n mean clm;
var ldl;
where ldl > 0;
run;

*Triglycerides;
proc means data=six n mean clm;
var trig;
where trig > 0;
run;

/************************/
/* Year 5 */
/************************/
data ef; set yr8.ef;
keep id_new dob sex ;
run;

proc sort data=ef;
by id_new;
run;

proc sort data=ryb;
by id_new;
run;

data year5; merge
ef (in=a)
ryb (in=b);
by id_new;
```

```
if a=b;
run;

*Yearly follow up form and anth form;
data foa; set yr8.foa;
where visit = 60;
run;

data anth; set yr8.anth;
where visit = 60;
run;

proc sort data=foa;
by id_new;
run;

proc sort data=anth;
by id_new;
run;

data y5; merge
foa (in=a)
anth (in=b);
by id_new;
if a=b;
run;

*Combining y5 and year5;
data yr5; merge
y5 (in=a)
year5 (in=b);
by id_new;
if a=b;
run;

*Age at year 5;
proc freq data=yr5;
tables dob foadat;
run;

data yr5_one; set yr5;
age_5 = (foadat - dob)/365.25;
run;

proc means data=yr5_one n min max mean clm;
var age_5;
run;
```

```
*Sex;
proc freq data=yr5;
tables sex;
run;

*BMI;
proc freq data=yr5;
tables hgt1 wgt1;
run;

data bmi; set yr5;
if hgt1 < 0 then delete;
if wgt1 < 0 then delete;
run;

data bmi2; set bmi;
hgt1m = hgt1/100;
bmi = wgt1/(hgt1m*hgt1m);
run;

proc means data=bmi2 n mean clm;
var bmi;
run;

*Weight;
proc means data=yr5 n mean clm;
var wgt1;
where wgt1 > 0;
run;

proc means data=yr5 n mean clm;
var wgt2;
where wgt2 > 0;
run;

proc means data=yr5 n mean clm;
var wgt3;
where wgt3 > 0;
run;

*SBP;
proc means data=yr5 n mean clm;
var sbp1;
where sbp1 > 0 ;
run;

*DBP;
proc means data=yr5 n mean clm;
```

```
var dbp1;
where dbp1 > 0;
run;

*A1C;
data lab; set yr8.central_lab;
where visit = 60;
run;

proc contents data=lab;
run;

proc sort data=lab;
by id_new;
run;

data lab1; merge
lab (in=a)
yr5 (in=b);
by id_new;
if a=b;
run;

proc means data=lab1 n mean clm;
var lab_HBA1C;
where lab_HBA1C > 0;
run;

*HDL cholesterol;
proc means data=lab1 n mean clm;
var lab_hdl;
where lab_hdl > 0;
run;

*Non-LDL;
data lab2; set lab1;
non_ldl = lab_chol - lab_hdl;
run;

proc means data=lab2 n mean clm;
var non_ldl;
where lab_CHOL > 0 ;
run;

*LDL;
proc means data=lab1 n mean clm;
var lab_LDL;
where lab_ldl > 0;
```

```
run;

*Triglycerides;
proc means data=lab1 n mean clm;
var lab_TG_NET;
where lab_TG_NET >0;
run;
```