Dataset Integrity Check for the Longitudinal Assessment of Bariatric Surgery 2 (LABS-2) Data Files

Prepared by Allyson Mateja
IMS, Inc.
3901 Calverton Blvd, Suite 200 Calverton, MD, 20705
September 2, 2016

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 Longitudinal Assessment of Bariatric Surgery (LABS) Consortium is a multicenter observational cohort study at ten US hospitals in six geographically diverse clinical centers. LABS had 3 phases; the second phase (LABS-2) focused on longer-term safety, outcomes, and durability of health changes. The major priorities for LABS-2 were to determine weight, medical, surgical, and behavioral outcomes, including incidence and remission of comorbid conditions, and to evaluate patient, procedure, and other characteristics that were associated with these outcomes. LABS-2 recruited adults undergoing first-time bariatric procedures between 2006 and 2009, and were followed up until September 2012. Participants completed research assessments prior to surgery and at 6 months, 12 months, and then annually after surgery. Three years after Roux-en-Y gastric bypass (RYGB) or laparoscopic adjustable gastric banding (LAGB), percent weight change from baseline and the percentage of participants with diabetes achieving hemoglobin A1c levels less than 6.5%or fasting plasma glucose values less than 126mg/dL without pharmacologic therapy were assessed. Dyslipidemia and hypertension resolution at 3 years were also examined. It was found that among participants with severe obesity, there was substantial weight loss 3 years after bariatric surgery, with the majority experiencing maximum weight change during the first year. However, there was variability in the amount and trajectories of weight loss and in diabetes, blood pressure, and lipid outcomes.

3 Archived Datasets

The SAS data files, as provided by the Data Coordinating Center (DCC), are located in the "Data" folder in the data package. For this replication, variables were taken from the "sq", "po1", "age_at_surgery", "calcvar", "post2", "mort", "rcab", and "sbp" data files.

4 Statistical Methods

Analyses were performed to duplicate results for the data published by Courcoulas, et al. in JAMA in 2013 [1]. To verify the integrity of the datasets, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], <u>Baseline Characteristics of the LABS-2 Cohort by Procedure</u>, Table A lists the variables that can be used in the replication. Table B compares the results calculated from the archived data file to the results published in Table 1. The results of the replication are almost an exact match.

For Table 2 in the publication [1], <u>Observed and Weighted Remission and Incident Rates 3 Years After Bariatric Surgery by Procedure</u>, Table C lists the variables that can be used in the replication. Table D compares the results calculated from the archived data file to the results published in Table 2. The results of the replication are almost an exact match.

For Table 3 in the publication [1], <u>Deaths and Subsequent Bariatric Surgery Procedures Within 3 Years of Initial Bariatric Surgery</u>, Table E lists the variables that can be used in the replication. Table F compares the results calculated from the archived data file to the results published in Table 3. The results of the replication are almost an exact match.

6 Conclusions

The NIDDK repository is confident that the LABS-2 data files to be distributed are a true copy of the manuscript data.

7 References

[1] Courcoulas, A.P., Christian, N.J., Belle, S.H., Berk, P.D., Flum, D.R., Garcia, L., Horlick, M., Kalarchian, M.A., King, W.C., Mitchell, J.E., Patterson, E.J., Pender, J.R., Pomp, A., Pories, W.J., Thirlby, R.C., Yanovski, S.Z., Wolfe, B.M. Weight Change and Health Outcomes at 3 Years After Bariatric Surgery Among Individuals with Severe Obesity. JAMA (2013); 310(22): 2416-2425.

Table A: Variables used to replicate Table 1: Baseline Characteristics of the LABS-2 Cohort by Procedure

Table Variable	dataset.variable
Procedure	sq.surg
Age	age_at_surgery.age_s
Weight	rcab.wgt, po1.wgt
BMI	rcab.wgt, po1.wgt, po1.hgtft, po1.hgtin
Sex	po1.sex
Race	po1.racew, po1.raceb, po1.racea, po1.racei, po1.raceh, po1.raceo
Ethnicity	po1.ethn
Diabetes	calcvar.dm2_p
Dyslipidemia	calcvar.dyslipid_p
Hyperlipidemia	calcvar.hlipid_p
Low HDL	calcvar.lowHDL
High triglycerides	calcvar.highTG_p
Hypertension	calcvar.htn_p

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

Characteristic	Overall Manuscript (N = 2458)	Overall DSIC (N = 2458)	Difference (N = 0)	Roux-en-Y Gastric Bypass Manuscript (n = 1738)	Roux-en-Y Gastric Bypass (n = 1738)	Difference (n = 0)
Age, median (IQR), y	46 (37-54)	46 (37-54)	0 (0-0)	45 (37-54)	45 (37-54)	0 (0-0)
Range, y	18-78	18-78	0-0	19-75	19-75	0-0
Weight, median (IQR), kg	129 (115-147)	128 (115-147)	1 (0-0)	131 (116-150)	130 (116-150)	1 (0-0)
Range, kg	75-290	75-289	0-1	75-240	75-239	0-1
BMI, median (IQR)	45.9 (41.7-51.5)	45.9 (41.8-51.4)	0 (0.1-0.1)	46.6 (42.4-51.9)	46.5 (42.4-51.8)	0.1 (0-0.1)
Range	33.0-94.3	32.8-94.1	0.2-0.2	33.7-81.0	33.7-81.0	0-0
Sex, No. (%)						
Female	1931 (78.6)	1931 (78.6)	0 (0)	1389 (79.9)	1389 (79.9)	0 (0)
Male	527 (21.4)	527 (21.4)	0 (0)	349 (20.1)	349 (20.1)	0 (0)
Race, No. (%)						
White	2102 (86.4)	2102 (86.4)	0 (0)	1463 (85.1)	1463 (85.1)	0 (0)
Black	256 (10.5)	256 (10.5)	0 (0)	196 (11.4)	196 (11.4)	0 (0)
Other	75 (3.1)	75 (3.1)	0 (0)	61 (3.5)	61 (3.6)	0 (0.1)
Ethnicity, No. (%)						
Hispanic	119 (4.8)	119 (4.9)	0 (0.1)	85 (4.9)	85 (4.9)	0 (0)
Non-Hispanic	2337 (95.2)	2337 (95.2)	0 (0)	1652 (95.1)	1652 (95.1)	0 (0)
Diabetes, No. (%)	774 (33.4)	774 (33.5)	0 (0.1)	583 (35.4)	583 (35.4)	0 (0)
Dyslipidemia, No. (%)	1252 (63.4)	1254 (63.5)	2 (0.1)	901 (64.4)	903 (64.4)	2 (0)
Hyperlipidemia, No. (%)	725 (36.6)	725 (36.5)	0 (0.1)	515 (36.7)	515 (36.6)	0 (0.1)
Low HDL, No. (%)	883 (37.5)	883 (37.5)	0 (0)	648 (38.8)	648 (38.8)	0 (0)
High triglycerides, No. (%)	462 (22.9)	463 (23.0)	1 (0.1)	339 (23.8)	340 (23.8)	1 (0)
Hypertension, No. (%)	1601 (67.5)	1601 (67.5)	0 (0)	1159 (68.9)	1159 (68.9)	0 (0)

Characteristic	Laparoscopic Adjustable Gastric Band Manuscript (n = 610)	Laparoscopic Adjustable Gastric Band DSIC (n = 610)	Difference (n = 0)	Sleeve Gastrectomy Manuscript (n = 59)	Sleeve Gastrectomy DSIC (n = 59)	Difference (n = 0)
Age, median (IQR), y	48 (37-56)	48 (37-56)	0 (0-0)	48 (36-55)	48 (35-55)	0 (1-0)
Range, y	18-78	18-78	0-0	21-73	21-73	0-0
Weight, median (IQR), kg	123 (111-139)	122 (111-139)	1 (0-0)	158 (134-180)	158 (132-181)	0 (2-1)
Range, kg	85-246	85-245	0-1	82-290	82-289	0-1
BMI, median (IQR)	43.9 (40.4-48.0)	43.7 (40.4-48.0)	0.2 (0-0)	57.7 (46.8-64.1)	57.7 (46.6-64.2)	0 (0.2-0.1)
Range	33.0-87.3	32.8-87.2	0.2-0.1	35.5-94.3	35.6-94.1	0.1-0.2
Sex, No. (%)						
Female	465 (76.2)	465 (76.2)	0 (0)	39 (66.1)	39 (66.1)	0 (0)
Male	145 (23.8)	145 (23.8)	0 (0)	20 (33.9)	20 (33.9)	0 (0)
Race, No. (%)						
White	543 (89.6)	543 (89.6)	0 (0)	47 (82.5)	47 (82.5)	0 (0)
Black	51 (8.4)	51 (8.4)	0 (0)	8 (14.0)	8 (14.0)	0 (0)
Other	12 (2.0)	12 (2.0)	0 (0)	2 (3.5)	2 (3.5)	0 (0)
Ethnicity, No. (%)						
Hispanic	26 (4.3)	26 (4.3)	0 (0)	6 (10.2)	6 (10.2)	0 (0)
Non-Hispanic	583 (95.7)	583 (95.7)	0 (0)	53 (89.8)	53 (89.8)	0 (0)
Diabetes, No. (%)	164 (28.8)	164 (28.8)	0 (0)	15 (28.8)	15 (28.9)	0 (0.1)
Dyslipidemia, No. (%)	291 (60.9)	291 (60.9)	0 (0)	33 (64.7)	33 (64.7)	0 (0)
Hyperlipidemia, No. (%)	177 (36.7)	177 (36.7)	0 (0)	22 (43.1)	22 (43.1)	0 (0)
Low HDL, No. (%)	194 (33.3)	194 (33.3)	0 (0)	21 (37.5)	21 (37.5)	0 (0)
High triglycerides, No. (%)	103 (21.1)	103 (21.1)	0 (0)	12 (21.8)	12 (21.8)	0 (0)
Hypertension, No. (%)	367 (62.7)	367 (62.7)	0 (0)	44 (80.0)	44 (80.0)	0 (0)

Characteristic	BPDS Manuscript (n = 19)	BPDS DSIC (n = 19)	Difference (n = 0)	Banded Gastric Bypass Manuscript (n = 32)	Banded Gastric Bypass DSIC (n = 32)	Difference (n = 0)
Age, median (IQR), y	39 (35-46)	39 (33-47)	0 (2-1)	48 (40-54)	48 (40-54)	0 (0-0)
Range, y	26-60	26-60	0-0	21-69	21-69	0-0
Weight, median (IQR), kg	136 (123-151)	136 (120-153)	0 (3-2)	136 (116-157)	136 (116-157)	0 (0-0)
Range, kg	110-192	110-192	0-0	92-227	97-227	0-0
BMI, median (IQR)	50.0 (44.9-52.3)	49.9 (44.2-52.6)	0.1 (0.7-0.3)	49.2 (42.5-54.1)	49.1 (42.3-54.3)	0.1 (0.2-0.2)
Range	37.9-62.6	38.0-62.5	0.1-0.1	36.2-76.0	36.0-75.9	0.2-0.1
Sex, No. (%)						
Female	14 (73.7)	14 (73.7)	0 (0)	24 (75.0)	24 (75.0)	0 (0)
Male	5 (26.3)	5 (26.3)	0 (0)	8 (25.0)	8 (25.0)	0 (0)
Race, No. (%)						
White	18 (94.7)	18 (94.7)	0 (0)	31 (100.0)	31 (100.0)	0 (0)
Black	1 (5.3)	1 (5.3)	0 (0)	0 (0.0)	0 (0.0)	0 (0)
Other	0 (0.0)	0 (0.0)	0 (0)	0 (0.0)	0 (0.0)	0 (0)
Ethnicity, No. (%)						
Hispanic	0 (0.0)	0 (0.0)	0 (0)	2 (6.2)	2 (6.3)	0 (0.1)
Non-Hispanic	19 (100.0)	19 (100.0)	0 (0)	30 (93.8)	30 (93.8)	0 (0)
Diabetes, No. (%)	7 (38.9)	7 (38.9)	0 (0)	5 (17.2)	5 (17.2)	0 (0)
Dyslipidemia, No. (%)	9 (52.9)	9 (52.9)	0 (0)	18 (64.3)	18 (64.3)	0 (0)
Hyperlipidemia, No. (%)	5 (29.4)	5 (29.4)	0 (0)	6 (21.4)	6 (21.4)	0 (0)
Low HDL, No. (%)	5 (27.8)	5 (27.8)	0 (0)	15 (50.0)	15 (50.0)	0 (0)
High triglycerides, No. (%)	5 (27.8)	5 (27.8)	0 (0)	3 (10.7)	3 (10.7)	0 (0)
Hypertension, No. (%)	10 (52.6)	10 (52.6)	0 (0)	21 (67.7)	21 (67.7)	0 (0)

Table C: Variables used to replicate Table 2: Observed and Weighted Remission and Incident Rates 3 Years After Bariatric Surgery by Procedure

Table Variable	dataset.variable
Procedure	sq.surg
Diabetes	calcvar.dm2_p
Dyslipidemia	calcvar.dyslipid_p
Hyperlipidemia	calcvar.hlipid_p
Low HDL	calcvar.lowHDL
High triglycerides	calcvar.highTG_p
Hypertension	calcvar.htn_p

Table D: Comparison of values computed in integrity check to reference article Table 2 values

	Roux-en-Y Gastric Bypass Manuscript (n = 1691)	Roux-en-Y Gastric Bypass DSIC (n = 1692)	Difference (n = 1)	Laparoscopic Adjustable Gastric Band Manuscript (n = 588)	Laparoscpoic Adjustable Gastric Band DSIC (n = 587)	Difference (n = 1)
	Observed No./Total No. (%)			Observed No./Total No. (%)		
Diabetes						
Remission	216/320 (67.5)	216/320 (67.5)	0/0 (0)	28/98 (28.6)	28/98 (28.6)	0/0 (0)
Incidence	5/560 (0.9)	5/562 (0.9)	0/2 (0)	8/247 (3.2)	8/247 (3.2)	0/0 (0)
Dyslipidemia						
Remission	237/383 (61.9)	237/383 (61.9)	0/0 (0)	39/144 (27.1)	39/144 (27.1)	0/0 (0)
Incidence	7/221 (3.2)	7/222 (3.2)	0/1 (0)	15/94 (16.0)	15/94 (16.0)	0/0 (0)
Hyperlipidemia						
Remission	151/253 (59.7)	151/253 (59.7)	0/0 (0)	22/97 (22.7)	22/97 (22.7)	0/0 (0)
Incidence	9/353 (2.5)	9/354 (2.5)	0/1 (0)	21/143 (14.7)	21/143 (14.7)	0/0 (0)
Low HDL						
Remission	292/341 (85.6)	293/342 (85.7)	1/1 (0.1)	76/113 (67.3)	76/113 (67.3)	0/0 (0)
Incidence	9/616 (1.5)	9/616 (1.5)	0/0 (0)	10/266 (3.8)	10/266 (3.8)	0/0 (0)
High triglycerides						
Remission	139/162 (85.8)	139/162 (85.8)	0/0 (0)	36/58 (62.1)	36/58 (62.1)	0/0 (0)
Incidence	8/495 (1.6)	8/496 (1.6)	0/1 (0)	14/206 (6.8)	14/206 (6.8)	0/0 (0)
Hypertension						
Remission*	269/705 (38.2)	290/740 (39.2)	21/35 (1.0)	43/247 (17.4)	51/256 (19.9)	8/9 (2.5)
Incidence*	39/309 (12.6)	40/331 (12.1)	1/22 (0.5)	27/149 (18.1)	28/164 (17.1)	1/15 (1.0)

*The variables in the data package that were used to calculate hypertension were corrected after the manuscript had been published. Therefore, the large differences between the DSIC and Manuscript values are expected.

Table E: Variables used to replicate Table 3: Deaths and Subsequent Bariatric Surgery Procedures Within 3 Years of Initial Bariatric Surgery

Table Variable	dataset.variable
Procedure	sq.surg
Death within 30 days of Surgery	post2.postdie
Cause of Death	mort.mortd, mort.mortds
Subsequent bariatric surgery procedures	sbp.procprio, sq.surgdat sq.surg, sq.age_s, sbp.nbaripro
Revision	sq.op_revis
Reversal	sq.op_rever
Band replacement	sq.op_rever, sq.op_revis
Port revision	sq.surg, sbp.portprob
Other revision	sq.surg, sq.op_rever, sq.op_revis, sbp.portprob
Band removal	sq.surg, sq.op_rever, sq.op_revis, sbp.portprob
Revision to another bariatric procedure	sq.surg, sq.op_rever, sq.op_revis, sbp.portprob

Table F: Comparison of values computed in integrity check to reference article Table 3 values

	No. of Participants Manuscript	No. of Participants DSIC	Difference
Roux-en-Y Gastric Bypass (n = 1738)			
Deaths	16	16	0
Within 30 days of surgery	3	3	0
Sepsis	1	1	0
Cardiovascular disease	1	1	0
Pulmonary embolism	1	1	0
More than 30 days after surgery	13	13	0
Bowel obstruction	1	1	0
Sepsis	1	1	0
Respiratory failure	1	1	0
Cardiovascular disease	3	3	0
Suicide/substance abuse	2	2	0
Cancer	1	1	0
Indeterminate after adjudication	4	4	0
Subsequent bariatric surgery procedures	4	4	0
Revision	2	2	0
Reversal	2	2	0
Laparoscopic Adjustable Gastric Band (n=610)			
Deaths	5	5	0
Within 30 days of surgery	0	0	0
More than 30 days after surgery	5	5	0
Organ failure	2	2	0
Respiratory failure	1	1	0
Cancer	1	1	0
Indeterminate after adjudication	1	1	0
Subsequent bariatric surgery procedures	77	78	1
Band replacement	7	7	0
Port revision	19	19	0
Other revision	10	11	1
Band removal	21	21	0
Revision to another bariatric procedure	20	20	0

Attachment A: SAS Code

```
**** LABS-2 DSIC;
**** Programmer: Allyson Mateja;
**** Date: July 5, 2016;

title1 "%sysfunc(getoption(sysin))";
title2 " ";

libname labs2 "/prj/niddk/ims_analysis/LABS/private_orig_data/Longitudinal Assessment of Bariatric Surgery (LABS-2)/SAS Database";

options nofmterr;

proc format;
      value sexf 1 = 'M'
                 2 = 'F';

      value surgf 1 = 'Roux-en-Y Gastric Bypass'
                  3 = 'BPDS'
                  4 = 'Laparoscopic Adjustable Gastric Band'
                  5 = 'Sleeve Gastrectomy'
                  7 = 'Other'
                  8 = 'Banded Gastric Bypass';

      value ethnf 0  = 'Hispanic'
                  1  = 'Non-Hispanic';

      value mortf 0       = 'Indeterminate after adjudication'
                  2,3     = 'Sepsis'
                  4       = 'Pulmonary embolism'
                  8       = 'Bowel obstruction'
                  11      = 'Respiratory failure'
                  13,16   = 'Suicide/substance abuse'
                  15      = 'Cancer'
                  6,7,17  = 'Cardiovascular disease'
                  18      = 'Organ failure';

/*data ae;             set labs2.ae;                 run;
data agb;              set labs2.agb;                run;
data agbp;            set labs2.agbp;                run;*/
data age_at_surgery;   set labs2.age_at_surgery;     run;
/*data ascf;           set labs2.ascf;               run;
data bb;              set labs2.bb;                  run;
data bdi;             set labs2.bdi;                 run;
data bf;              set labs2.bf;                  run;
data bio_status;      set labs2.bio_status;          run;
data bpds;            set labs2.bpds;                run;
data bs;              set labs2.bs;                  run;
data bu;              set labs2.bu;                  run;*/
data calcvar;         set labs2.calcvar;             run;
/*data cd;             set labs2.cd;                 run;
data cdf;             set labs2.cdf;                 run;
data cdfm;            set labs2.cdfm;                run;
```

```
data central_lab_results; set labs2.central_lab_results; run;
data clab_status;        set labs2.clab_status;         run;
data dib;                set labs2.dib;                 run;
data dif;                set labs2.dif;                 run;*/
data ds;                 set labs2.ds;                  run;
/*data ec;                 set labs2.ec;                  run;*/
data ef;                 set labs2.ef;                  run;
/*data ef_nocons;          set labs2.ef_nocons;           run;
data eq5d;               set labs2.eq5d;                run;
data ess;                set labs2.ess;                 run;
data fo6;                set labs2.fo6;                 run;
data gen_status;         set labs2.gen_status;          run;
data gs;                 set labs2.gs;                  run;
data gsrs;               set labs2.gsrs;                run;*/
data in2;                set labs2.in2;                 run;
/*data in2_excluded;       set labs2.in2_excluded;        run;
data isel;               set labs2.isel;                run;
data iw;                 set labs2.iw;                  run;
data leak;               set labs2.leak;                run;
data ll;                 set labs2.ll;                  run;*/
data mab;                set labs2.mab;                 run;
/*data maf;                 set labs2.maf;                 run;
data maq;                set labs2.maq;                 run;
data med;                set labs2.med;                 run;
data medrx;              set labs2.medrx;               run;*/
data mort;               set labs2.mort;                run;
/*data mvf;                 set labs2.mvf;                 run;
data mwf;                set labs2.mwf;                 run;
data ndi;                set labs2.ndi;                 run;
data niv;                set labs2.niv;                 run;
data off;                set labs2.off;                 run;
data path;               set labs2.path;                run;
data petsb;              set labs2.petsb;               run;
data petsf;              set labs2.petsf;               run;*/
data po1;                set labs2.po1;                 run;
data post2;              set labs2.post2;               run;
/*data preg_max_by_id;      set labs2.preg_max_by_id;      run;*/
data pu2;                set labs2.pu2;                 run;
data rcab;               set labs2.rcab;                run;
/*data rcaf;                set labs2.rcaf;                run;
data rhb;                set labs2.rhb;                 run;
data rhf;                set labs2.rhf;                 run;
data rhp;                set labs2.rhp;                 run;
data rsf;                set labs2.rsf;                 run;
data rsi;                set labs2.rsi;                 run;
data ryb;                set labs2.ryb;                 run;*/
data sbp;                set labs2.sbp;                 run;
/*data sbq;                 set labs2.sbq;                 run;
data sbqf;               set labs2.sbqf;                run;
data sf36;               set labs2.sf36;                run;
data sfb;                set labs2.sfb;                 run;
data sff;                set labs2.sff;                 run;
data short;              set labs2.short;               run;
data sl;                 set labs2.sl;                  run;
data sls;                set labs2.sls;                 run;
data smab;               set labs2.smab;                run;
```

14

```
data smaf;                set labs2.smaf;            run;*/
data sq;                  set labs2.sq;              run;
/*data sw_minute;            set labs2.sw_minute;         run;
data sw_summary;          set labs2.sw_summary;      run;
data uevnt;               set labs2.uevnt;           run;
data uib;                 set labs2.uib;             run;
data uif;                 set labs2.uif;             run;
data vtype;               set labs2.vtype;           run;
data wef;                 set labs2.wef;             run;
data wgt;                 set labs2.wgt;             run;
data whq;                 set labs2.whq;             run;
data wpai;                set labs2.wpai;            run;*/

proc sort data = ef nodupkey;
       by id;

proc sort data = sq;
       by id surgdat;

/*data sq;
       set sq;
       by id;
       if last.id then output;*/

proc sort data = po1 nodupkey;
       by id;

proc sort data = pu2 ;
       by id surgdat;

proc sort data = age_at_surgery;
       by id surgdate;

proc sort data = mab nodupkey;
       by id;

proc freq data = calcvar;
       tables visit;

data baseline_calcvar;
       set calcvar;
       if visit = 1;

data year3_calcvar;
       set calcvar;
       if visit = 36;

proc sort data = baseline_calcvar nodupkey;
       by id;

proc contents data = sq;

proc sort data = rcab nodupkey;
       by id;

data po_updates;
```

15

```
        merge sq (in=val1)
              pu2 (keep=id surgdat wgt)
              age_at_surgery (keep=id surgdate primary age_s canceled rename = (surgdate = surgdat));
        by id surgdat;
        if val1 then output po_updates;

data surgeries;
        merge ef  (in=val1 keep=id)
              po1 (keep = id po1dat age_c wgt sex racew raceb racea racei raceh raceo ethn hgtft hgtin)
              po_updates (in=val2 rename = (wgt = wgt_upd))
              baseline_calcvar (keep=id dm2_p dyslipid_p hlipid_p lowHDL highTG_p htn_p)
              rcab (in=val1 keep=id wgt rename = (wgt=rcab_wgt));
        by id;
        if racew=1 and raceb=0 and racea=0 and racei=0 and raceh=0 and raceo=0 then race = 'W';
        else if racew=0 and raceb=1 and racea=0 and racei=0 and raceh=0 and raceo=0 then race = 'B';
        else if racew in (-3, -4) and raceb in (-3,-4) and racea in (-3,-4) and racei in (-3,-4) and raceh in (-3,-4) and raceo in (-3,-4) then race = '';
        else race = 'O';
        if val1 and val2 then output surgeries;

proc freq data = surgeries;
        tables surgno /list missing;

data surgeries;
        set surgeries;
        if surgno = 0 ;

proc freq data = surgeries;
        tables surg*primary /list missing;

data surgeries secondary_surgery;
        set surgeries;
        if primary = 1 then output surgeries;
        else output secondary_surgery;

data surgeries;
        set surgeries;
        surg_time = surgdat - po1dat;
        if wgt_upd ne . and surg_time > 30 then wgt = wgt_upd;
        wgt_kg = round(wgt*0.453592,1);
        rcab_wgt_kg = round(rcab_wgt*0.453592,1);
        if rcab_wgt_kg < 0 then rcab_wgt_kg = wgt_kg;
        height_total = (hgtft + (hgtin/12))*0.3048;
        bmi = round(rcab_wgt_kg/(height_total**2), 0.1);

proc freq data = surgeries;
        tables surg;
        format surg surgf.;
        title3 'Table 1 - Procedure';

proc sort data = surgeries;
        by surg;

proc means data = surgeries n median p25 p75 min max;
        var age_s;
        class surg;
        types () surg;
```

16

```
        format surg surgf.;
        title3 'Table 1 - Age';

proc means data = surgeries n median p25 p75 min max;
        var rcab_wgt_kg;
        class surg;
        types () surg;
        format surg surgf.;
        title3 'Table 1 - Weight';

proc means data = surgeries n median p25 p75 min max;
        var bmi;
        class surg;
        types () surg;
        format surg surgf.;
        title3 'Table 1 - BMI';

proc freq data = surgeries;
        tables sex;
        format sex sexf.;
        title3 'Table 1 - Sex';

proc freq data = surgeries;
        tables sex;
        by surg;
        format sex sexf. surg surgf.;

proc freq data = surgeries;
        tables race ;
        title3 'Table 1 - Race';

proc freq data = surgeries;
        tables race;
        by surg;
        format surg surgf.;

proc freq data = surgeries;
        tables ethn;
        where ethn ne -3;
        format ethn ethnf.;
        title3 'Table 1 - Ethnicity';

proc freq data = surgeries;
        tables ethn;
        where ethn ne -3;
        by surg;
        format ethn ethnf. surg surgf.;

proc freq data = surgeries;
        tables dm2_p;
        title3 'Table 1 - Diabetes';

proc freq data = surgeries;
        tables dm2_p;
        by surg;
        format surg surgf.;
```

17

```
proc freq data = surgeries;
      tables dyslipid_p;
      title3 'Table 1 - Dyslipidemia';

proc freq data = surgeries;
      tables dyslipid_p;
      by surg;
      format surg surgf.;

proc freq data = surgeries;
      tables hlipid_p;
      title3 'Table 1 - Hyperlipidemia';

proc freq data = surgeries;
      tables hlipid_p;
      by surg;
      format surg surgf.;

proc freq data = surgeries;
      tables lowHDL;
      title3 'Table 1 - Low HDL';

proc freq data = surgeries;
      tables lowHDL;
      by surg;
      format surg surgf.;

proc freq data = surgeries;
      tables highTG_p;
      title3 'Table 1 - High Triglycerides';

proc freq data = surgeries;
      tables highTG_p;
      by surg;
      format surg surgf.;

proc freq data = surgeries;
      tables htn_p;
      title3 'Table 1 - Hypertension';

proc freq data = surgeries;
      tables htn_p;
      by surg;
      format surg surgf.;

proc sort data = year3_calcvar;
      by id;

proc sort data = surgeries;
      by id;

proc sort data = in2 nodupkey;
      by id;

proc sort data = post2 nodupkey;
```

```
        by id surgdat;

proc sort data = mort nodupkey;
        by id;

proc sort data = mort;
        by id surgdat;

data table3;
        merge surgeries (in=val1 keep=id surg surgdat age_c age_s op_rever op_revis surgs)
              post2     (keep=id surgdat postdie)
              mort      (keep=id surgdat mortd mortds mortdat age_d);
        by id surgdat;
        death_time = age_d-age_c;
        year_time = floor((mortdat-surgdat)/365.25);
        if year_time > 3 and death_time = 3 then death_time = 4;
        if mortds in ('COMPLICATIONS OF HYPERTENSIVE CV DISEASE', 'FATAL ARRHYTHMIA/CARDIOMYOPATHY') then mortd = 17;
        if mortds in ('LIVER/HEPATIC FAILURE', 'MSOF') then mortd = 18;
        if val1 and surg in (1,4) then output table3;

proc sort data = table3 nodupkey;
        by id;

data table2;
        merge year3_calcvar (in=val1 keep=id dm2_p dyslipid_p hlipid_p lowHDL highTG_p htn_p rename = (dm2_p = dm2_p_yr3 dyslipid_p = dyslipid_p_yr3 hlipid_p =
hlipid_p_yr3
                            lowHDL = lowHDL_yr3 highTG_p = highTG_p_yr3 htn_p = htn_p_yr3))
              surgeries     (in=val2 keep=id surg age_s surgdat dm2_p dyslipid_p hlipid_p lowHDL highTG_p htn_p rename = (dm2_p = dm2_p_bl dyslipid_p =
dyslipid_p_bl hlipid_p = hlipid_p_bl
                            lowHDL = lowHDL_bl highTG_p = highTG_p_bl htn_p = htn_p_bl))
              in2           (keep=id age_i in2dat)
              table3        (keep=id death_time postdie mortd);
        by id;
        inactive_time = (age_i-age_s);
        inactive_time2 = floor((in2dat-surgdat)/365.25);
        if mortd = . or death_time > 3 then death = 0;
        else if (mortd ne . and 0 <= death_time <= 3) then death = 1;
        if dm2_p_bl = 0 and dm2_p_yr3 = 1 then diabetes_incidence = 1;
        else if dm2_p_bl = 0 and dm2_p_yr3 = 0 then diabetes_incidence = 0;
        if dm2_p_bl = 1 and dm2_p_yr3 = 0 then diabetes_remission = 1;
        else if dm2_p_bl = 1 and dm2_p_yr3 = 1 then diabetes_remission = 0;
        if dyslipid_p_bl = 0 and dyslipid_p_yr3 = 1 then dyslipidemia_incidence = 1;
        else if dyslipid_p_bl = 0 and dyslipid_p_yr3 = 0 then dyslipidemia_incidence = 0;
        if dyslipid_p_bl = 1 and dyslipid_p_yr3 = 0 then dyslipidemia_remission = 1;
        else if dyslipid_p_bl = 1 and dyslipid_p_yr3 = 1 then dyslipidemia_remission = 0;
        if hlipid_p_bl = 0 and hlipid_p_yr3 = 1 then hyperlipidemia_incidence = 1;
        else if hlipid_p_bl = 0 and hlipid_p_yr3 = 0 then hyperlipidemia_incidence = 0;
        if hlipid_p_bl = 1 and hlipid_p_yr3 = 0 then hyperlipidemia_remission = 1;
        else if hlipid_p_bl = 1 and hlipid_p_yr3 = 1 then hyperlipidemia_remission = 0;
        if lowHDL_bl = 0 and lowHDL_yr3 = 1 then lowHDL_incidence = 1;
        else if lowHDL_bl = 0 and lowHDL_yr3 = 0 then lowHDL_incidence = 0;
        if lowHDL_bl = 1 and lowHDL_yr3 = 0 then lowHDL_remission = 1;
        else if lowHDL_bl = 1 and lowHDL_yr3 = 1 then lowHDL_remission = 0;
        if highTG_p_bl = 0 and highTG_p_yr3 = 1 then highTG_incidence = 1;
        else if highTG_p_bl = 0 and highTG_p_yr3 = 0 then highTG_incidence = 0;
        if highTG_p_bl = 1 and highTG_p_yr3 = 0 then highTG_remission = 1;
```

19

```
            else if highTG_p_bl = 1 and highTG_p_yr3 = 1 then highTG_remission = 0;
            if htn_p_bl = 0 and htn_p_yr3 = 1 then htn_incidence = 1;
            else if htn_p_bl = 0 and htn_p_yr3 = 0 then htn_incidence = 0;
            if htn_p_bl = 1 and htn_p_yr3 = 0 then htn_remission = 1;
            else if htn_p_bl = 1 and htn_p_yr3 = 1 then htn_remission = 0;
            if val1 and val2 and ((inactive_time = . or inactive_time > 3) or (inactive_time2 = . or inactive_time2 > 3)) then output table2;

data table2;
            set table2;
            if surg in (1,4) and death = 0;

proc freq data = table2;
            tables surg /list missing;
            format surg surgf.;
            title3 'Table 2 - Procedures';

proc sort data = table2;
            by surg;

proc freq data = table2;
            tables diabetes_remission diabetes_incidence /list;
            by surg;
            format surg surgf.;
            title3 'Table 2 - Diabetes';

proc freq data = table2;
            tables dyslipidemia_remission dyslipidemia_incidence /list;
            by surg;
            format surg surgf.;
            title3 'Table 2 - Dyslipidemia';

proc freq data = table2;
            tables hyperlipidemia_remission hyperlipidemia_incidence /list;
            by surg;
            format surg surgf.;
            title3 'Table 2 - Hyperlipidemia';

proc freq data = table2;
            tables lowHDL_remission lowHDL_incidence /list;
            by surg;
            format surg surgf.;
            title3 'Table 2 - Low HDL';

proc freq data = table2;
            tables highTG_remission highTG_incidence /list;
            by surg;
            format surg surgf.;
            title3 'Table 2 - High triglycerides';

proc freq data = table2;
            tables htn_remission htn_incidence /list ;
            by surg;
            format surg surgf.;
            title3 'Table 2 - Hypertension';

proc freq data = table3;
```

```
        tables surg;
        format surg surgf.;
        title3 'Table 3 - Procedures';

proc sort data = table3;
        by surg;

proc freq data = table3;
        tables postdie;
        by surg;
        format surg surgf.;

proc freq data = table3;
        tables mortd /list;
        where postdie = 1;
        by surg;
        format surg surgf. mortd mortf.;

proc freq data = table3;
        tables mortd /list ;
        where postdie ne 1 and (death_time <= 3);
        by surg;
        format surg surgf. mortd mortf.;

proc sort data = sq;
        by id surgdat;

proc sort data = secondary_surgery;
        by id surgdat;

proc sort data = post2;
        by id surgdat;

proc sort data = sbp;
        by id surgdat;

data subsequent_surgeries;
        merge sq                  (in=val1 keep=id surg surgdat op_rever op_revis surgs)
              secondary_surgery   (in=val2 keep=id surgdat age_s primary canceled)
              post2               (keep=id surgdat postdie)
              sbp                 (keep=id surgdat revbari nbaripro portprob procprio);
        by id surgdat;
        if val1 and val2 and postdie = 0 and procprio = 1 then output subsequent_surgeries;

proc sort data = surgeries;
        by id;

proc sort data = table3;
        by id;

data subsequent_surgeries;
        merge subsequent_surgeries (in=val1)
              surgeries            (keep=id surgdat age_s surg rename = (surgdat = orig_surgdat age_s = orig_age_s surg = orig_surg));
        by id;
        surg_time = ((surgdat - orig_surgdat)/365.25);
        surg_time2 = age_s - orig_age_s;
```

21

```
        if val1 and (surg_time <= 3 and surg_time2 <= 3) and orig_surg in (1,4) then output;

data subsequent_surgeries;
        length sub_surg $50.;
        set subsequent_surgeries;
        if orig_surg = 1 then do;
                if op_rever = 1 then sub_surg = 'Reversal';
                if op_revis = 1 then sub_surg = 'Revision';
        end;
        if orig_surg = 4 then do;
                if surg = 7 then sub_surg = 'Port revision';
                if surg in (1,5) then sub_surg = 'Revision to another bariatric procedure';
                if surg = 4 and op_rever = 1 and op_revis = 1 then sub_surg = 'Band replacement';
                if surg = 4 and ((op_rever = 1 and op_revis = 0) or portprob = 1) then sub_surg = 'Band removal';
                if surg = 4 and op_rever = 0 and op_revis = 1 and portprob = 0 then sub_surg = 'Other revision';
                if surg = 4 and nbaripro = 1 then sub_surg = 'Ignore';
        end;
        if sub_surg ne 'Ignore';


proc sort data = subsequent_surgeries;
        by orig_surg;

proc freq data = subsequent_surgeries;
        tables sub_surg /list missing;
        by orig_surg;
        format orig_surg surgf.;
```