Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Inge TH, Courcoulas AP, Jenkins TM, et al. Five-year outcomes of gastric bypass in adolescents as compared with adults. N Engl J Med 2019;380:2136-45. DOI: 10.1056/NEJMoa1813909

Table of Contents

APPENDIX A. Design and Methods

History and features of LABS and Teen-LABS study design

To facilitate and accelerate data collection and research in adult bariatric surgery, in 2003 the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) established the Longitudinal Assessment of Bariatric Surgery (LABS) consortium.[1] LABS was designed to assess the short and longer term outcomes of bariatric surgery and to evaluate its role in the understanding and treatment of obesity and its complications in adults (age ≥ 18 years old). The consortium used standardized measures and procedures to systematically follow consecutive cases of metabolic and bariatric surgery as performed as part of standard medical care at 6 clinical centers involving 11 hospitals in the U.S. LABS-1 was an initial 30 day safety study.[2] LABS-2, an extension of the LABS-1study, consented participants for a comprehensive, longer term assessment of safety and efficacy of surgery over up to 7 years. LABS-2 also evaluated the effect of bariatric surgery on psychosocial and behavioral outcomes.[3,4] A Data Coordinating Center was created to facilitate the project, and a central laboratory (Northwest Lipid Laboratory, Seattle, WA) was selected.

Teen-LABS, an ancillary study to LABS, was proposed in 2006 as a prospective observational cohort study, collecting coordinated clinical, epidemiological, and behavioral data in adolescent bariatric surgical patients.[5] The design of Teen-LABS was patterned on both LABS-1 and LABS-2, using similar research methodology and data collection instruments. The LABS and Teen-LABS studies were planned in a coordinated and collaborative manner to afford the first opportunity to understand broad ranging outcomes of gastric bypass surgery in these two distinct populations with duration of obesity as the differentiating factor. Teen-LABS investigators designed and with assistance of the LABS consortium, deployed a weigth history questionnaire[6] for completion by LABS-2 study participants. These data allowed for selection of the 398 LABS-2 adult participants who declared a personal history of obesity to age 18 and underwent gastric bypass during adulthood. This comparison to the Teen-LABS participants who underwent gastric bypass as adolescents permits a realistic estimate of the risks and benefits of bariatric surgery in adolescent years and adulthood and will lead to a better understanding of the plasticity of important medical and psychosocial obesity-related comorbidities. Teen-LABS was originally designed and powered to compare the outcomes of gastric bypass between adolescents and adults who carried obesity forward from adolescence enrolled in the LABS-2 study – the parent study to Teen-LABS.

The Teen-LABS study was conducted as a cooperative agreement and funded by the National Institute of Diabetes and Digestive and Kidney Diseases with a grant to University of Colorado, Denver (UM1 DK072493) PI: Thomas Inge, MD, PhD and the University of Cincinnati (UM1 DK095710) PI: Changchun Xie, PhD and Todd Jenkins, PhD, MPH. We gratefully acknowledge the significant contributions made by the Teen-LABS Consortium as well as our parent study LABS Consortium (U01 DK066557).

These current analyses address the main adolescent-adult comparative aims that were proposed initially in the grant funding the Teen-LABS study. For the purposes of this combined analysis, LABS-2 participants were used, and are referred to in the manuscript simply as “LABS participants.”

Informed Consent

Written informed consent was obtained from all subjects according to the guidelines established by the Institutional Review Board each participating site. All participants (or their legally authorized

representative) were provided with a consent form before entering the study or undergoing any study-specific procedures. An investigator or study coordinator reviewed the consent form and answered questions. Four specific consent/permission forms to address the special circumstances of this study:

- Informed Parental Permission for minor participants, signed by at least one parent or legal guardian; coupled with
    - Assent for minor participants
- Informed Consent for use in re-consenting adolescent participants who became 18 years old during course of the study
- Informed Consent for participants enrolling at age ≥18 years

Baseline and follow-up data were collected within 30 days before operation, at the time of discharge from the hospital, at 30 days, at 6 months (window of 3-9 months), 1 year (window of 9-18 months), 2 year (window of ±6 months), 3 year (window of ±6 months), 4 year (window of ±6 months), and 5 year (window of ±6 months) postoperative research visits.

Height and Weight

Height was measured using the same device for pre and postoperative measurements. At each center, a calibrated wall-mounted stadiometer was used. For home visits, a stadiometer was shipped to the field examiner and calibrated prior to the visit. Height measurements were also made in triplicate.

Preoperative measurement of weight was obtained at the time of the enrollment visit and on the same Tanita scale (Tanita model TBF-310, Tokyo, Japan) at each clinical visit. Tanita scales were shipped to the field examiner for home visits and calibrated prior to the visit. Measurements were obtained with patients in light clothing and without shoes. Weight measurements were obtained in triplicate and recorded to the nearest 100 grams.

Weight change was reported in kilograms and as the percentage of change from baseline. During in-person assessments, weight was measured on a standard scale (Tanita Body Composition Analyzer, model TBF-310). If this per-protocol weight was not obtained, weights were measured on a non-study scale, and if neither was available, a participant's self-reported weight was used. Measurements were obtained with patients in light clothing and without shoes. Weight measurements were obtained in triplicate and recorded to the nearest 100 grams. Differences between in-person and self-reported weights when both were available in these cohorts were small (1.1 kg or less).[7,8] Weights of women in their second or third trimester of pregnancy or up to 6 months post-partum were excluded from analyses.

Comorbidity prevalence, remission, and incidence definitions

Comorbidity prevalence was calculated as the number of individuals meeting case criteria for the condition, divided by the number of individuals with evaluable data who were eligible to have that condition at baseline or follow-up. Comorbidity remission was calculated as the percentage of subjects without the condition at post-operative time points, among those with who had the condition at baseline and had evaluable data at follow-up. Comorbidity incidence was calculated as the percentage of subjects with the condition at post-operative time points, among those who did not have the condition at baseline.

Below we have provided the criteria used in defining the presence of a case of a comorbid condition. We have generally used standard conventions, based on laboratory abnormalities and medication use as

applicable. These research definitions may differ somewhat from those used in diagnosing conditions in a clinical setting. In a clinical setting, often multiple observations over time are required (for instance for a diagnosis of hypertension), or in some conditions, specialized testing may be needed to aid a clinician in confirming that patient meets diagnostic criteria. These factors should be taken into consideration when considering prevalence and remission data.

Diabetes mellitus (DM). DM at baseline was defined by study investigators taking into consideration patient self-report of prior diagnosis as well as prior medical records from referring physician, use of medications for DM, baseline HbA1c of ≥6.5%, or fasting glucose of ≥126 mg/dL, or oral glucose tolerance results in the prior 6 months. Participants reporting having polycystic ovary syndrome who did not meet laboratory criteria for DM and were not taking a DM medication other than metformin were not considered to have diabetes. Participants who were on metformin at baseline for weight management or for insulin resistance with no other indication of a prior diagnosis of DM documented and no laboratory findings consistent with the diagnosis of DM were not considered to have DM.

Remission of DM: Unless otherwise noted, remission of DM was defined as no use of medication for DM, and HbA1c < 6.5% (if HbA1c was not available, remission also required FBG <126mg/dL). In instances where specified laboratory and/or medication use data were unavailable, subject-reported declarations of presence of or absence of diabetes were used (as documented prospectively at each study visit on "medical assessment baseline" and "medical assessment follow-up" case report forms). Subject-reported information was required for defining cases in 3.6% of Teen-LABS and 21% of LABS participants.

Hypertension. Blood pressure (BP) was measured at the time of the study visit and use of medications for control of BP was recorded on medication use form (MED). For this analysis, hypertension was defined in a manner consistent with that used to clinically define hypertension: use of BP medications or systolic BP>140 mmHg or diastolic BP > 90 mmHg.

Remission of hypertension (HTN): Unless otherwise noted, remission of HTN required that no medications for BP were being used and systolic BP < 140 mmHg and diastolic BP < 90 mmHg. Specifically, the data for this variable were obtained as described below:

- Systolic and diastolic BP were measured using a Welch Allyn Spot Vital Signs monitor 4200B. For home visits, a monitor was shipped to the field examiner.
- Measurement of BP was done with appropriately sized cuff and after the patient has been seated quietly, with feet flat on the floor, in an erect but comfortable posture for at least five minutes, and for at least thirty minutes since the patient has smoked or consumed caffeine-containing beverages.

Hypertriglyceridemia. Hypertriglyceridemia was defined for those <21 years of age as fasting triglycerides (TG) ≥130 mg/dL, or for those 21 and older, hypertriglyceridemia was defined as fasting TG ≥200 mg/dL. Remission of hypertriglyceridemia was defined as TG < 130 mg/dL for those < 21 years of age, and TG <200 mg/dL for those ≥ 21 years of age, in an individual who met criteria for hypertriglyceridemia at baseline.

Low high-density lipoprotein cholesterol. High density lipoprotein cholesterol (HDL-C) was considered abnormally low if the measured HDL-C was < 40 mg/dL for males or was <50 mg/dL for females.

Remission of low HDL was defined as HDL-C ≥40 mg/dL (males) or HDL-C ≥50 mg/dL (females), in an individual who had previously met criteria for low HDL-C at baseline.

Remission of dyslipidemia: If <21 years of age, at follow-up, remission of dyslipidemia was defined as TG <130 mg/dL, and LDL-C <130 mg/dL, and HDL-C ≥40 mg/dL, and no use of LLM. If age was ≥ 21 years, resolution of dyslipidemia was defined as TG <200 mg/dL, and LDL-C <160 mg/dL, and HDL-C ≥40 mg/dL (males) or HDL-C ≥50 mg/dL (females), and no use of LLM.
Specifically, the data for this variable were obtained as follows:

- Central laboratory measured triglyceride, LDL cholesterol, HDL cholesterol at baseline and follow-up;
- LLM assessment was derived during analysis from Comorbidity Assessment-Baseline (CAB) or follow up (CAF) form, Question 5 – selection equals: "treatment with single medication for dyslipidemia" or "treatment with two or more medications for dyslipidemia";
- Medications (MED) form, subject-reported use of any antilipidemic medication.

Adjudication Process
To accurately and objectively assess the risks of bariatric surgery, investigators and NIH considered it critical that deaths among study participants in Teen-LABS and LABS be clearly classified as related to the surgical intervention or to other causes unrelated to the surgical intervention. The Adjudication Committee for Teen-LABS is a group of professionals with expertise in surgery, pediatrics, obesity, and mental health who are not study investigators but who have volunteered to review and classify etiology and relatedness to bariatric surgery of adjudicatable events. LABS investigators served on an analogous LABS adjudication committee with expertise pertinent to adults. Clinical records and study data were stripped of participant and site identifiers and were shared with Adjudication Committee members, for review and classification of events.

Laboratory Analyses
All laboratory assays were performed by the Northwest Lipid Metabolism and Diabetes Research Laboratories (Seattle, WA). Reference ranges for serum chemistries as identified by Northwest Lipid Laboratory are shown in Table S5.

Statistical Methods
Cohort-specific categorical measures are presented using frequencies and percentages and compared using chi-square or Fisher's exact tests. Continuous variables were summarized using means with standard deviations or medians with intra-quartile range; t-test or Wilcoxon Rank-Sum tests used to compare by cohort. All statistical analyses were conducted using SAS v9.4; all reported p-values were two-sided and considered statistically significant when less than 0.05. No adjustments were made for multiple comparisons.

Repeated measures analyses were performed using mixed models with a subject-level random intercept term. Sex, baseline household income, and baseline education were forced into each model, as these characteristics were associated with missing follow-up visits (Table S6). Linear mixed models were used to compare percent weight change from baseline by study cohort. The following variables were considered for inclusion in the final model: race, ethnicity, baseline body weight. Modeled percentages and 95% confidence intervals were calculated by cohort and study visit. Weight values from female participants in their second or third trimester of pregnancy and up to six months postpartum were omitted

from analyses.

Poisson mixed modeling with robust error variance was used to compare prevalence (comorbidities and micronutrient outcomes) and remission by study cohort, with sex, baseline household income, and baseline education were forced into each model. Modeled means and 95% confidence intervals were calculated by cohort and study visit. The following variables were considered for inclusion in the final models:

Diabetes prevalence: race, ethnicity, baseline diabetes medication use, body weight, baseline HbA1c, baseline duration of diabetes;

Diabetes remission: race, ethnicity, baseline diabetes medication use, baseline body weight, percent weight change from baseline, baseline HbA1c, baseline duration of diabetes;

Hypertension prevalence: race, ethnicity, body weight;

Hypertension remission: race, ethnicity, baseline blood pressure medication use, baseline body weight, percent weight change from baseline, baseline systolic blood pressure, baseline diastolic blood pressure;

Low HDL prevalence: race, ethnicity, body weight; Low HDL remission: race, ethnicity, baseline body weight, percent weight change from baseline, baseline HDL; High triglycerides prevalence: race, ethnicity, body weight; High triglycerides remission: race, ethnicity, baseline body weight, percent weight change from baseline, baseline triglycerides;

Low ferritin prevalence: race, ethnicity, body weight;

Low vitamin B12 prevalence: race, ethnicity, body weight;

Low vitamin D prevalence: race, ethnicity, body weight.

Two-year micronutrient analyses were conducted on all Teen-LABS participants and a subset of LABS subjects (N=179). The micronutrients analyzed were measured by the central laboratory within days of collection in the Teen-LABS subjects but were not measured in real time for LABS participants. For LABS participants, a request for specimens that had been stored in the NIH biorepository was made and only for a subset of the group eventually selected to be in the comparison cohort for this manuscript. These specimens were shipped to the same central laboratory used for the Teen-LABS samples.

Poisson modeling was used to calculate death and intra-abdominal 5-year event rates by cohort. Incidence rates and 95% confidence intervals were calculated by study cohort and intra-abdominal event category. Rates are expressed per 500-person years (i.e., 100 patients followed for 5 years). Sex and race were forced into each model, as these characteristics were associated with missing intra-abdominal event data. Incidence rate ratios (IRR) and 95% confidence intervals were calculated to compare total intra-abdominal event rates between adolescents and adults. The following variables were considered for inclusion in the final model that generated the IRR value: sex, race, ethnicity, baseline weight, baseline body fat percentage, baseline diabetes status, baseline hypertension status, and method of RYGB procedure (open, laparoscopic).

A sensitivity analysis was performed with the diabetes remission outcome model due to missing values with the duration of diabetes model covariate term. Among those with diabetes at baseline, the self-reported duration of diabetes data was available for 65% of adolescents and 41% of adults – thereby greatly reducing the number of observations available for the modeling analysis. Although not found to be a significant term and dropped from the final model in the results reported in the main report, we conducted a sub-analysis where duration of diabetes was forced into the final model. Just as with the final model in the main report, duration of diabetes was not a significant term in the model (p=0.37). Also, the relationship between adolescents and adults for diabetes remission was similar (Risk Ratio: 1.16; 95% CI: 1.02, 1.32; p=0.029 compared to risk ratio: 1.27; 95% CI: 1.03, 1.57; p=0.03 reported in the main findings).

Appendix B. Missing Data Sensitivity Analyses

Sensitivity analyses were performed to evaluate the missing at random (MAR) assumption by:

(1) Comparing baseline characteristics between those with and without 5 year weight data;
(2) Weight data gathered beyond 5 years;
(3) Conducting sensitivity analyses using pattern mixture modeling.

(1) Baseline Characteristics Comparison

The tables below present select baseline characteristics for those with and without 5 year weight data, by TL (Table S7) and LABS (Table S8) cohorts. For TL subjects, no significant differences were noted across any baseline characteristics. Among LABS subjects, those missing 5 year weight data were more likely to be female and white, and have a greater weight at baseline.

(2) Weight data gathered beyond 5 years

To address the question of whether the missing data from those participants who missed their 5 year study visit could be an important source of bias, we recognized that since both study populations were still participating in longitudinal follow-up visits, we had the opportunity to examine weight values beyond the 5 year visit in both groups. Comparison of the weight data beyond 5 years for those who missed and those who didn't miss the 5 year visit inform us about whether the individuals who missed their 5 year visit have similar or different weight outcomes than those who did not miss. Therefore, utilizing body weight data collected at the 6 and 7 year study visits, we were able to evaluate the longer term body weight endpoint between those did and didn't complete their 5 year visits. Using linear mixed modeling, we calculated estimates and 95% confidence intervals for those with and without 5 year body weight data. Figures S5 and S6 display these results for Teen-LABS (Figure S5) and LABS (Figure S6) cohorts. For Teen-LABS, the wide, overlapping confidence intervals point to similar estimates between groups (Figure S5). Based on the point estimates, those missing their 5 year visit appear to exhibit superior weight loss at 6 and 7 years. Thus, if there is bias introduced, that bias would be toward underestimating the effect of surgery on weight-related outcomes. Figure S6 for LABS participants demonstrates nearly identical 6 to 7 year weight change between those with and without 5 year body weight data, suggesting that the adult LABS participants missing at year 5 are also likely not a major source of bias for our outcomes.

(3) Sensitivity analyses using pattern mixture modeling

Sensitivity analyses using pattern-mixture models were performed to evaluate the missing at random (MAR) assumption. A total of 100 imputed data sets were created for use in the multivariable models (using similar methods as in the main report). SAS Proc Mixed and Proc MiAnalyze were used to generate all estimates from the multiply imputed datasets. Using this approach, imputed percent weight change values were adjusted by -5%, 0%, and +5% of what they would be if the data were MAR. Based on the findings above that utilized 6 and 7 year data, the imputation shift range of +/- 5% was anticipated to provide conservative estimates. The figures below plot the imputed estimates for -5% (blue), 0% (black), and +5% (orange) shifts by TL (Figure S7) and LABS (Figure S8) populations. For both TL and LABS, imputed estimates were similar, supporting the validity of the MAR assumption.

APPENDIX C. Supplementary Tables

Table S1: Baseline characteristics by Study Group

	Teen-LABS (N=161)	LABS (N=396)	p value
Age at Surgery, Mean (SD)	17.0 (1.52)	37.9 (7.04)	<0.001
Sex, % (n)			0.57
Female	78.3% (126)	76.0% (301)	
Male	21.7% (35)	24.0% (95)	
Race, % (n)			0.17
White	73.9% (119)	80.1% (317)	
Black	21.7% (35)	15.2% (60)	
Other	4.4% (7)	4.7% (19)	
Ethnicity, % (n)			0.08
Non-Hispanic	90.7% (146)	94.7% (375)	
Hispanic	9.3% (15)	5.3% (21)	
Body Weight (kg), mean (SD)	150.9 (30.32)	145.6 (28.51)	0.05
Body Mass Index, mean (SD)	53.7 (9.63)	50.6 (7.81)	<0.001

Table S2: Mean Change in Clinical Variables Over 5 years by Study Group

	Teen-LABS 5 year Mean Change From Baseline (95% CI)	**LABS** 5 year Mean Change From Baseline (95% CI)	p-value
Weight (kg)	-37.3 (-41.6,-33.0)	-40.2 (-43.1,-37.3)	0.18
Body Mass Index (kg/m^2)	-12.7 (-14.2,-11.2)	-13.8 (-14.8,-12.8)	0.16
Systolic blood pressure (mmHg)	-7.4 (-11.4,-3.3)	-5.4 (-8.5,-2.3)	0.82
Diastolic blood pressure (mmHg)	-5.3 (-8.1,-2.5)	-3.1 (-5.3,-1.0)	0.63
Glycated hemoglobin (%)	-0.48 (-0.63,-0.32)	-0.52 (-0.63,-0.41)	0.62
HDL cholesterol (mg/dL)	16.7 (13.5,19.8)	15.9 (13.6,18.2)	0.16
Non-HDL cholesterol (mg/dL)	-24.5 (-32.8,-16.3)	-29.7 (-35.5,-23.9)	0.20
LDL cholesterol (mg/dL)	-13.6 (-19.8,-7.4)	-10.6 (-14.8,-6.4)	0.32
Triglycerides (mg/dL)	-61.0 (-84.5,-37.4)	-57.5 (-74.2,-40.8)	0.76

Table S3: Modeled and Observed Comorbidities

	Teen-LABS				LABS			
	Baseline		5 years		Baseline		5 years	
	Observed†	Modeled¥	Observed†	Modeled¥	Observed†	Modeled¥	Observed†	Modeled¥
Diabetes								
No. with data	161	161	139	139	388	388	223	223
Prevalence	17.6 (28/159)	13.6 (9.3,20.0)	2.2 (3/138)	2.4 (0.8,6.7)	36.1 (140/388)	31.1 (26.7,36.2)	15.7 (35/223)	12.2 (8.8,16.9)
Remission	--	--	85.7 (12/14)	85.9 (70.0, 100.0)	--	--	55.4 (36/65)	53.0 (42.0,67.0)
Incidence	--	--	0.9 (1/114)	0.17 (0.02,1.14)	--	--	0.7 (1/145)	0.52 (0.18,1.50)
Hypertension								
No. with data	159	159	136	136	385	385	234	234
Prevalence	35.9 (57/159)	29.6 (22.1,39.6)	15.2 (19/125)	14.9 (9.0,24.6)	62.3 (240/385)	61.3 (56.3,66.7)	41.5 (93/224)	39.1 (33.4,45.8)
Remission	--	--	70.2 (33/47)	67.5 (51.8,88.0)	--	--	41.5 (56/135)	41.2 (33.3,51.2)
Incidence	--	--	6.6 (5/76)	6.9 (2.8,16.9)	--	--	13.6 (11/81)	11.3 (6.3,20.0)
Low HDL cholesterol								
No. with data	160	160	124	124	384	384	197	197
Prevalence	65.6 (105/160)	53.2 (33.2,85.1)	14.0 (16/114)	12.8 (5.9,27.9)	44.5 (171/384)	36.8 (20.1,67.3)	8.4 (16/190)	6.5 (2.9,14.6)
Remission	--	--	80.5 (62/77)	78.4 (67.6,90.9)	--	--	82.9 (63/76)	83.5 (75.5,92.4)
Incidence	--	--	2.7 (1/37)	0.12 (0.02,0.65)	--	--	2.6 (3/114)	0.05 (0.02,0.15)
Hypertriglyceridemia								
No. with data	160	160	124	124	379	379	187	187
Prevalence	42.5 (68/160)	36.4 (27.8,47.6)	5.3 (6/114)	5.9 (2.5,14.3)	31.4 (119/379)	30.2 (25.7,35.5)	13.3 (24/180)	12.2 (8.5,17.7)
Remission	--	--	90.6 (48/53)	80.6 (67.6,96.1)	--	--	70.7 (41/58)	69.0 (58.7,81.1)
Incidence	--	--	1.7 (1/60)	0.0 (0.0,0.0)	--	--	6.1 (7/114)	2.1 (0.7,6.4)

† Observed prevalence, remission, and incidence data are expressed as arithmetic proportion (observed cases/eligible cases)
¥ Modeled prevalence, remission, and incidence data are expressed as modeled proportion (95% CI)

Table S4: Modeled Prevalence of Micronutrient Deficiency over 2 Years

	Teen-LABS				LABS				p-value*
	Baseline		2 years		Baseline		2 years		
	Observed†	Modeled¥	Observed†	Modeled¥	Observed†	Modeled¥	Observed†	Modeled¥	
Low Ferritin	4/160	1.8% (0.5,7.2)	72/132	48.3% (37.2,62.7)	4/178	2.2% (0.8,5.8)	54/179	28.5% (22.5,36.2)	0.004
Low Vitamin B12	1/159	0	14/132	3.7% (1.4, 9.7)	4/178	0.9% (0.3,2.7)	17/179	3.5% (1.7,7.3)	0.91
Low Vitamin D	71/159	24.5% (17.3,34.6)	64/131	37.8% (28.2,50.8)	81/178	35.5% (26.9,46.8)	38/179	23.8% (17.5,32.4)	0.020

† Observed prevalence data are expressed as arithmetic proportion (observed cases/eligible cases)
¥ Modeled prevalence data are expressed as modeled proportion (95% CI)

* Group comparison at 2-year time point.

Table S5: Laboratory normal reference ranges

Analyte	Normal reference range
Vitamin B12	180-914 pg/mL
25-OH Vitamin D	20.1-50 ng/mL
Ferritin, females	10-180 µg/L
Ferritin, males	20-230 µg/L
Glucose (fasting)	<110 mg/dL, 110-125 mg/dL borderline
LDL	Optimal: <100 Near Optimal: 100-129 Borderline: 130-159 High Risk: 160-189 Very High: ≥190
HDL	Low: <40 Very High: ≥ 60
Triglyceride	Optimal: <150 Borderline: 150-199 High Risk: ≥200

Table S6: Baseline predictors of post-operative missed study visits

	Crude models			Final model		
	OR	95% CI	p-value	OR	95% CI	p-value
White race	1.40	0.88, 2.26	0.16			
Female	0.59	0.40, 0.86	0.006	0.52	0.35, 0.78	0.001
Non-Hispanic	1.22	0.81, 1.83	0.35			
BMI	1.00	0.99, 1.01	0.99			
Type II diabetes	1.05	0.85, 1.30	0.66			
Dyslipidemia	1.03	0.81, 1.30	0.84			
Hypertension	1.19	0.98, 1.44	0.09			
Education			0.039			0.020
Less than High school	<REF>			<REF>		
HS graduate	0.52	0.32, 0.87		0.79	0.29, 2.15	
Some college	0.62	0.38, 0.99		0.64	0.24, 1.73	
College graduate	0.57	0.34, 0.95		0.62	0.21, 1.81	
Post-graduate work	0.78	0.46, 1.32		1.52	0.51, 4.51	
Household income			0.001			<0.001
Less than $25,000	<REF>			<REF>		
$25,000-49,999	1.52	1.15, 2.02		0.44	0.27, 0.71	
$50,000-74,999	1.83	1.37, 2.44		0.69	0.42, 1.13	
$75,000-99,000	1.26	0.85, 1.88		0.26	0.11, 0.63	
$100,000 or more	1.85	1.21, 2.82		0.25	0.09, 0.68	

Table S7: Baseline characteristics by 5 year weight availability – Teen-LABS

	5 yr weight available	5 yr weight not available	p-value
N	140	21	
Age at surgery, mean (SD)	16.9 (1.54)	17.4 (1.38)	0.17
Female, % (n)	79% (111)	71% (15)	0.41
White, % (n)	73% (102)	81% (17)	0.43
Non-Hispanic, % (n)	91% (127)	90% (19)	0.97
Weight (kg), mean (SD)	152 (30.70)	144 (27.33)	0.27
BMI, mean (SD)	54 (9.92)	52 (7.51)	0.49
Diabetes, % (n)	14% (19)	29% (6)	0.10
Hypertension, % (n)	37% (51)	29% (6)	0.46
Low HDL cholesterol, % (N)	65% (91)	67% (14)	0.91
High triglycerides, % (N)	41% (57)	52% (11)	0.33

Table S8: Baseline characteristics by 5 year weight availability – LABS

	5 yr weight available	5 yr weight not available	p-value
N	294	102	
Age at surgery, mean (SD)	38.1 (7.09)	37.1 (6.88)	0.18
Female, % (n)	81% (237)	63% (64)	<0.001
White, % (n)	78% (228)	87% (89)	0.04
Non-Hispanic, % (n)	95% (280)	93% (95)	0.42
Weight (kg), mean (SD)	144 (28.19)	151 (28.92)	0.03
BMI, mean (SD)	50 (7.85)	51 (7.72)	0.42
Diabetes, % (n)	31% (90)	33% (34)	0.73
Hypertension, % (n)	62% (176)	63% (64)	0.92
Low HDL cholesterol, % (N)	42% (120)	51% (51)	0.13
High triglycerides, % (N)	33% (92)	27% (27)	0.24

APPENDIX D. Supplementary Figures

Figure S1: CONSORT Flow Diagram

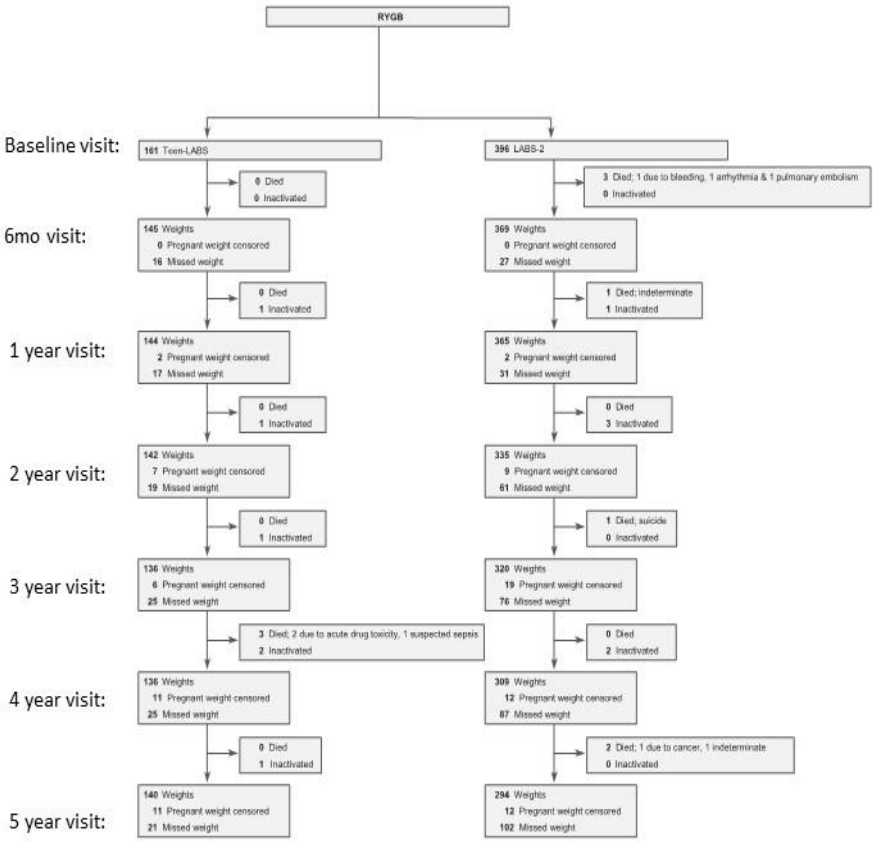


Figure S2: Diabetes Prevalence Over Time by Study Group

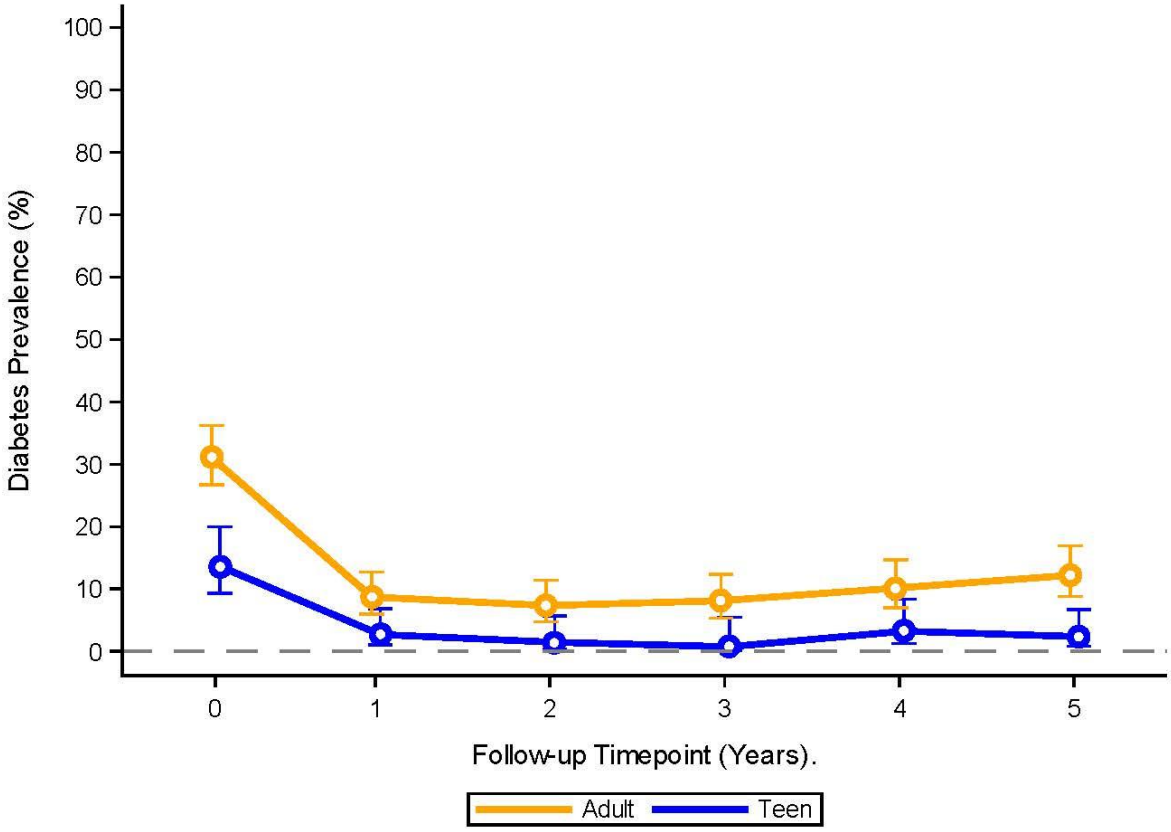


Figure S3: Risk Ratios for Comorbidity Prevalence Teen-LABS vs. LABS (REF)

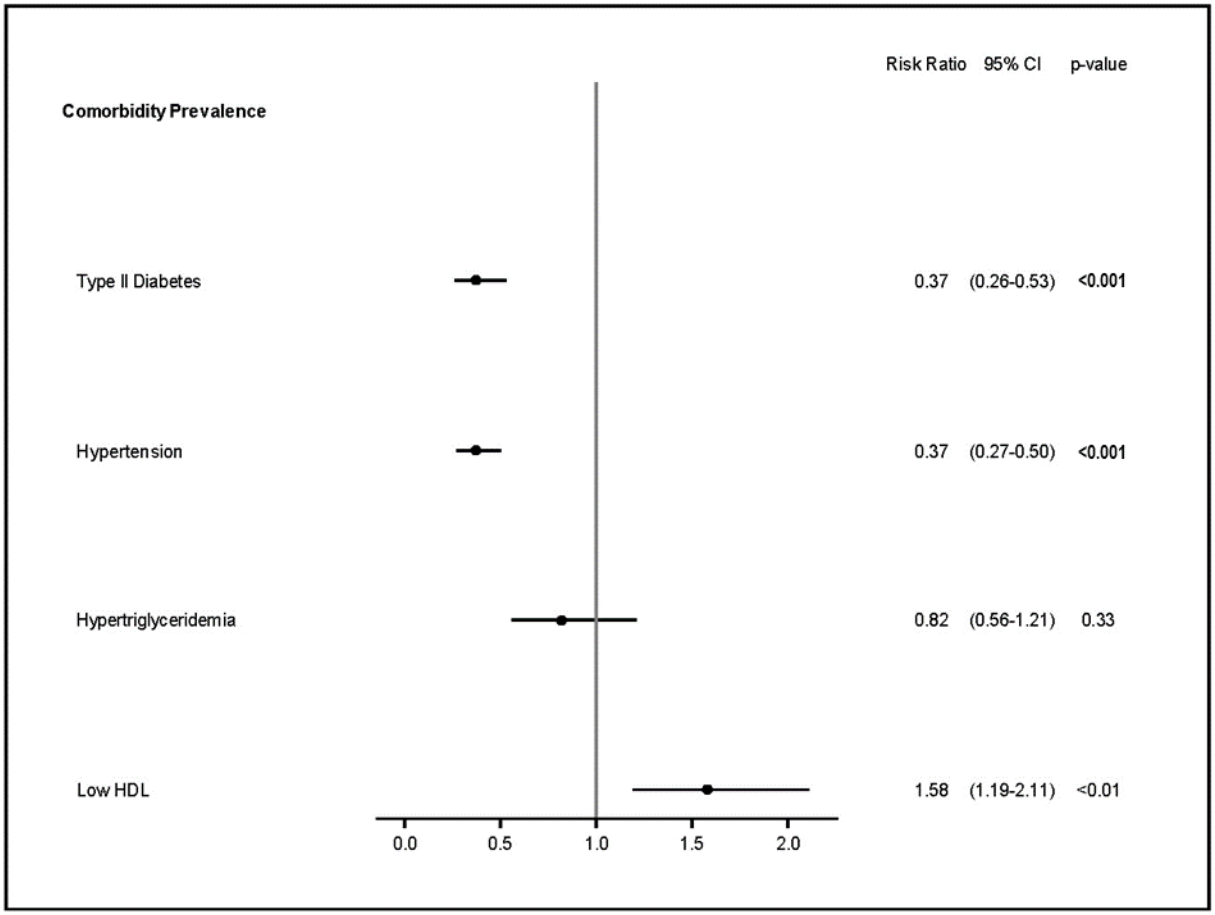


Figure S4: Hypertension Prevalence Over Time by Study Group

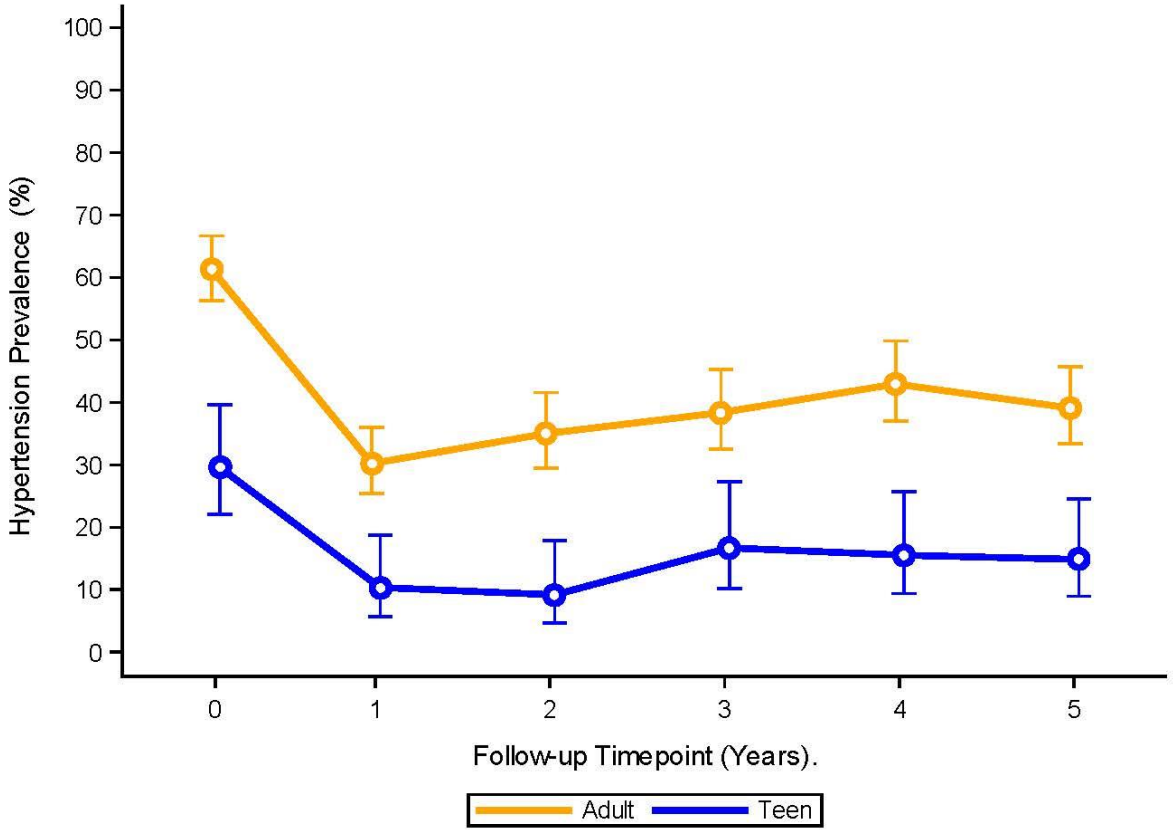


Figure S5: Weight Change by Availability of 5 year Weight data in Teen-LABS subjects

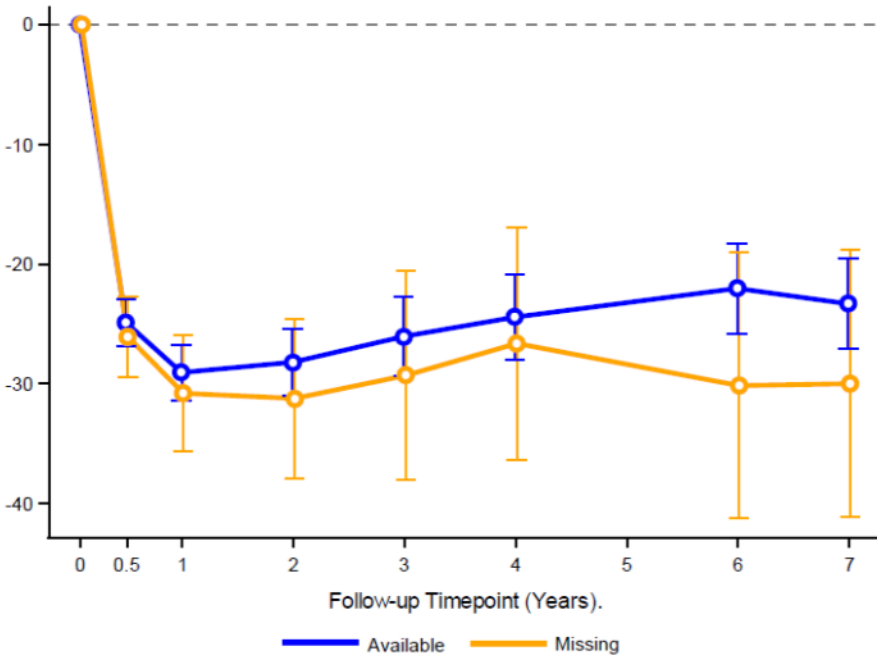


Figure S6: Weight Change by Availability of 5 year Weight data in LABS subjects

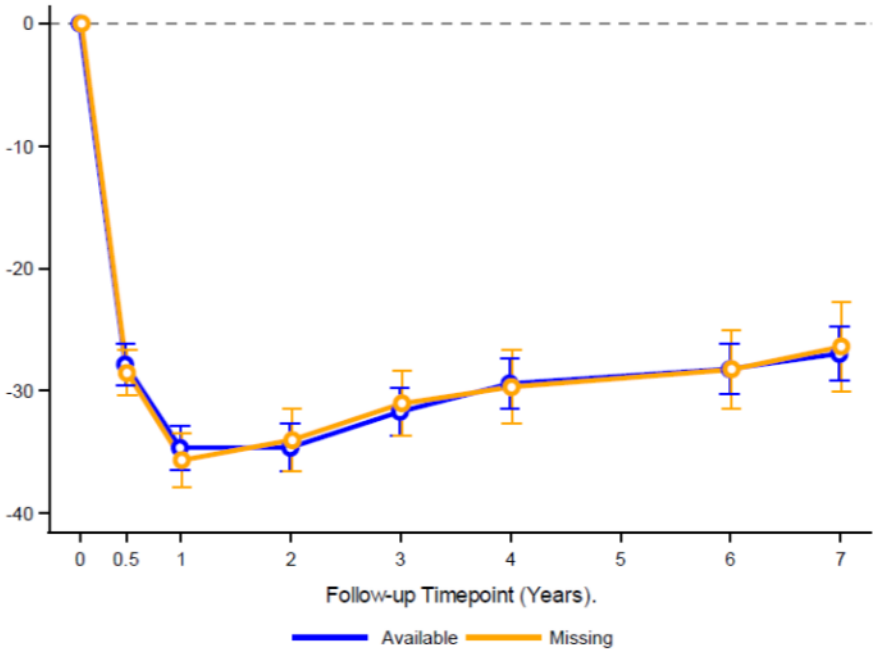


Figure S7: Weight Change by Imputation Shift in Teen-LABS subjects

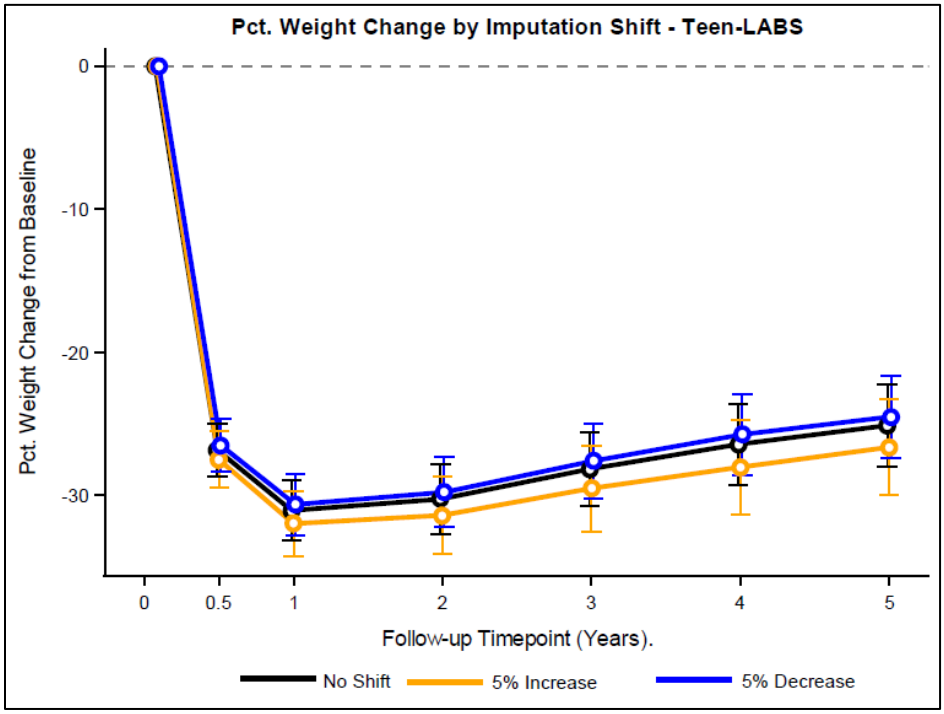
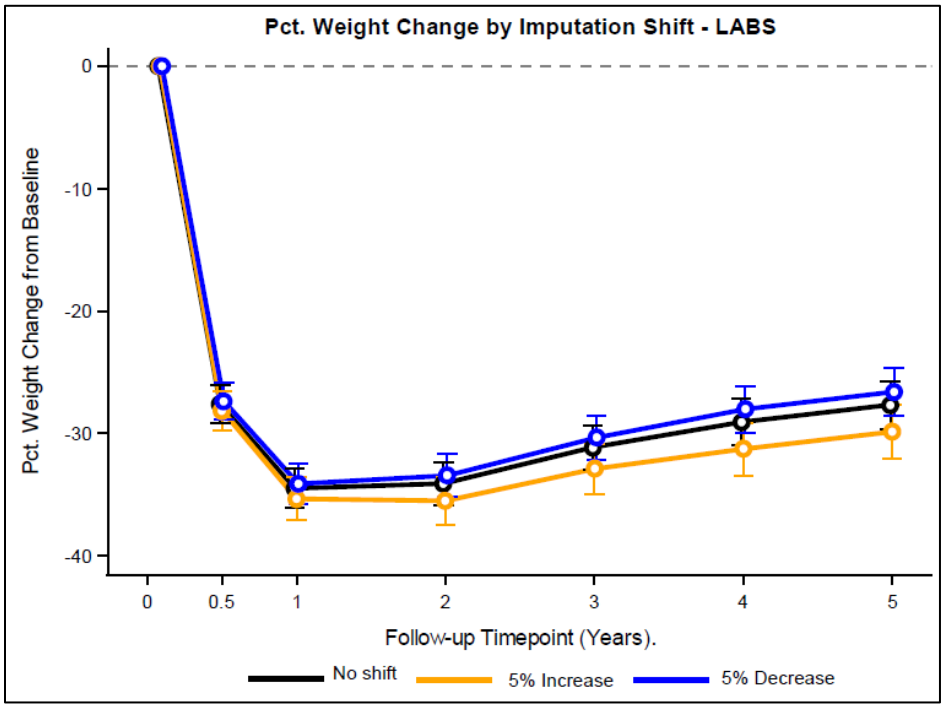


Figure S8: Weight Change by Imputation Shift in LABS subjects



Appendix E. Acknowledgements

The decision to publish the manuscript was made by all authors and the Teen-LABS Consortium and by collaborating authors in the LABS Consortium. The first author drafted the manuscript and all authors participated in critical reviews and editing. The DCC performed data analyses according to a plan approved by the steering committee and attests to the veracity and completeness of the data. All authors had full access to the data, critically reviewed and edited, vouch for the integrity and accuracy of the analyses, and made the decision to publish the manuscript. The sponsor collaborated in the study design, data analysis, and writing process but did not impose restrictions on the manuscript.

The authors would like to thank the participants and their families, as well as the talented clinical staff at each site for their dedication to the patient population represented in this study. The authors would also like to thank the central study coordinators, Rosie Miller, RN, CCRC and Jennifer Black, MSSA, LSW, and the study site coordinators (Jennifer Knox, BA, and Naila Merchant, Cincinnati Children's Hospital Medical Center; David Allen, MS, Texas Children's Hospital, Baylor College of Medicine; Beverly Haynes, BSN, Children's Hospital of Alabama, University of Alabama; Eleanor Shirley, MA, University of Pittsburgh Medical Center; Lauren Volz, BS, Nationwide Children's Hospital) for their extraordinary dedication and resourcefulness in conduct of the study. The consortium is also grateful for the special statistical expertise of Jane Khoury, PhD, and Ralph Buncher, ScD.

In addition, we gratefully acknowledge the significant contributions made by the LABS Consortium members, including Dr. Dave R. Flum, University of Washington; Dr. Steven H. Belle, University of Pittsburgh; Drs. Paul Berk and Alfons Pomp, New York Columbia-Presbyterian/Cornell University Medical Center; Drs. Walter Pories, and Kenneth MacDonald, East Carolina University; Dr. Anita Courcoulas, University of Pittsburgh; Dr. James Mitchell, Neuropsychiatric Research Institute; Dr. Bruce Wolfe, Oregon Health & Sciences University; and NIDDK project scientists, Drs. Carolyn Miles, Susan Yanovski, and Mary Horlick.

The authors also appreciate the important work done by the Adjudication Committees. Teen-LABS Adjudication Committee Members: Dan von Allmen, MD, Bruce Wolfe, MD, Chris Bolling, MD, Joy Collins, MD, David Wesson, MD, David Lanning, MD, PhD, Stephanie Walsh, MD, Nirav Desai, MD, Dana Rofey, PhD, Karin Price, PhD, Allen Browne, MD, Mark Wulkan, MD, Kirk Reichard, MD. LABS Adjudication Committee Members: Alfons Pomp, MD, FACS; Steven H. Belle, PhD, MScHyg; Marc Bessler, MD; E. Patchen Dellinger, MD; Laura Machado, MD; Emma Patterson, MD; Walter J. Pories, MD, FACS.

Finally, the authors would like to thank the following individuals for their helpful suggestions during preparation and editing of the manuscript: Aaron Kelly, PhD; Stephen Daniels, MD, PhD; Lisa McLeod, MD; Phil Zeitler, MD, PhD; Kristin Nadeau, MD, MS.

APPENDIX F. References

1. Belle SH, Berk PD, Courcoulas AP, et al. Safety and efficacy of bariatric surgery: Longitudinal Assessment of Bariatric Surgery. Surg Obes Relat Dis 2007;3:116-26.
2. Longitudinal Assessment of Bariatric Surgery C, Flum DR, Belle SH, et al. Perioperative safety in the longitudinal assessment of bariatric surgery. The New England journal of medicine 2009;361:445-54.
3. Courcoulas AP, Christian NJ, Belle SH, et al. Weight change and health outcomes at 3 years after bariatric surgery among individuals with severe obesity. Jama 2013;310:2416-25.
4. Courcoulas AP, King WC, Belle SH, et al. Seven-Year Weight Trajectories and Health Outcomes in the Longitudinal Assessment of Bariatric Surgery (LABS) Study. JAMA Surg 2018;153:427-34.
5. Inge TH, Zeller M, Harmon C, et al. Teen-Longitudinal Assessment of Bariatric Surgery: methodological features of the first prospective multicenter study of adolescent bariatric surgery. J Pediatr Surg 2007;42:1969-71.
6. Jenkins TM, Buncher CR, Akers R, et al. Validation of a weight history questionnaire to identify adolescent obesity. Obes Surg 2013;23:1404-12.
7. Jenkins TM, Boyce TW, Ralph Buncher C, et al. Accuracy of Self-Reported Weight Among Adolescent and Young Adults Following Bariatric Surgery. Obes Surg 2017;27:1529-32.
8. Christian NJ, King WC, Yanovski SZ, Courcoulas AP, Belle SH. Validity of self-reported weights following bariatric surgery. Jama 2013;310:2454-6.