Supplementary Online Content

Adams TD, Davidson LE, Litwin SE, et al. Health benefits of gastric bypass surgery after 6 years. *JAMA*. doi:10.1001/2012.jama.11164.

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

eMethods.

Participant Selection

At study initiation, gastric banding was not approved in the U.S. and vertical banded gastroplasty was no longer performed, limiting the surgical procedure studied to gastric bypass. Although all patients were from one high-volume bariatric surgical center, results from this certified center of excellence facility should be generalizable to other high-volume surgical centers.

The first control group was defined by severely obese participants also seeking gastric bypass surgery from the same surgical practice as those who had surgery. These participants were generally not approved for surgery because their current insurance policy did not include gastric bypass surgery as a covered procedure. One of the major insurers in Utah did not cover this procedure until very recently. While there was no attempt to match this control group with the surgery group, the same characteristics (with the exception of income level) that led to seeking surgery were expected to have been similar in the two groups. Because economic reasons could explain differences in insurance coverage or the ability to self-pay for the operation, we included education level, income level, and marital status in the definition of a propensity score to use as a covariate in the regression models to help control for the baseline group differences.

The second control group was defined by severely obese participants randomly selected from the Utah population using the Health Family Tree database[1,2] with only BMI≥35 kg/m^2 as a selection criterion. These severely obese participants were quite different at baseline, as intended, so as to provide a greater ability to generalize the findings of this study to all severely obese participants. Age, BMI, and gender distribution were the main variables that differed at baseline for this control group, so that these variables were also included in the propensity score definition or as covariates.

Exclusions from the study included prior weight-loss procedures, active cancer being treated within the last five years (except non-melanoma skin cancer), and myocardial infarction within the last 6 months.

Study Protocols

Most variables were obtained from either an examination at the University of Utah Cardiovascular Genetics outpatient clinic or at the Center for Clinical and Translational Science (CCTS) in the University of Utah hospital. Exams at both locations were performed by the same study staff, who were all trained and standardized in their measurement protocols. Same-brand weight scales and blood pressure devices were used at both locations. Study location was determined by the number of available beds each night at the CCTS and occasionally by participant preference not to stay overnight.

Biochemical variables were measured from a morning blood draw obtained after an overnight 12-hour fast, except for water. The blood samples from both locations were processed at Cardiovascular Genetics and sent to the University of Utah hospital-associated lab (ARUP) for metabolic panel, HbA1c, and complete blood count measurements. Blood pressures were measured in a sitting position after five minutes of rest alone in an exam room. Three blood pressures were obtained by a Dinamap automated blood pressure recorder (GE Healthcare, Tampa, FL) using a cuff size dictated by the mid-arm circumference. Because of the difficulty in measuring blood pressure in the severely obese, a blood pressure measurement that differed by more than 10 mmHg from the other two was repeated, and the average of three blood pressures that were within 10 mmHg of each other was used for analysis. Height was measured by a Harpenden anthropometer (Holtain, Ltd, Crymych, United Kingdom) to the nearest centimeter. Weight was measured by a Scaletronix scale (model 5100, Scaletronix Corporation, Wheaton, IL) that has an 800 pound capacity and accuracy of 0.1 kg. Waist circumference was measured at the umbilicus. After each participant had reclined quietly for at least five minutes, percent body fat was measured by bioelectrical impedance using an RJL Systems Analyzer (Quantum II, Clinton, MI). Because existing equations accompanying the RJL unit are not very accurate at the extremes of the BMI distribution, the reactance and resistance values obtained from the impedance test along with the participant's age, sex and body weight were entered into a percent body fat regression equation developed by our research team from a study of severely obese men and women who underwent hydrostatic weighing and bioelectrical impedance.[3,4] Glucose, insulin, HOMA-IR, HbA1c, triglycerides, and VLDL-C were log transformed for analyses using these variables as absolute amounts at any of the three exams. The two- and six-year changes of these variables were not log transformed.

Detailed medical histories were obtained when participants came to their baseline exam and were used to define prevalent disease at baseline. Current and past medications were itemized and the reason for taking each of the medications recorded. Identical procedures were used at each follow-up exam for those returning to our clinics. Medical records were searched at two of the main health care chains serving Utah (University of Utah and Intermountain Healthcare) for information on participants not returning for exams 2 or 3. In addition, consent by

participants to obtain medical records from their physicians was used to contact physician offices and to obtain such records. The completeness of the medical record data was supplemented when possible by phone interviews of the participants. This process of medical record abstraction from those records obtained from the two medical care systems or from physician offices provided a meaningful amount of additional clinical data. We note that the clinical data obtained was often not as complete as the total data acquired when the participant visited our outpatient clinic or the University hospital (CCTS). The effects of collecting incomplete data on these participants were assessed in our sensitivity analyses described below.

Diagnosis and procedural codes of all hospitalizations occurring in Utah are reported to the Utah Department of Health. All participants in our study (n = 1156) were matched with the Utah hospitalization data. This process specifically focused on matching any surgical or control patients who were hospitalized for reasons related to bariatric-specific surgical procedures or bariatric surgery-related complications using 138 CPT and ICD9 codes..

Medication Adjustment

Including a medication term in a regression equation does not adjust the medicated clinical or biochemical values to their unmedicated values, but merely removes the differences between the medicated and unmedicated participants. Therefore, to estimate the unmedicated levels, we used the earliest exam of the subset of participants who met the clinical definition of each endpoint (diabetes, hypertension, low HDL-C, high triglycerides, high LDL-C) but were not being treated by medication at that exam to calculate sex-specific mean levels. These mean levels from the unmedicated participants were assigned to all participants receiving medication for that indication, replacing their measured levels. The sex-averaged means of the variables assigned to the medicated participants were: systolic blood pressure, 152 mmHg; diastolic blood pressure, 96 mmHg; glucose, 160 mg/dl; insulin, 11.6 µU/ml; HbA1c, 6.8%; total cholesterol, 255 mg/dl; LDL-C, 169 mg/dl; HDL-C, 34 mg/dl; triglycerides, 279 mg/dl; and VLDL-C, 55 mg/dl. HOMA-IR was calculated as glucose X insulin / 22.5.

Sensitivity Analysis

To assess the influence of statistical assumptions of various models used for analysis, we performed a detailed sensitivity analysis of the data. The study results in Table 2 of the main paper show the medication- and propensity score-adjusted group six-year changes. We verified that the individual covariates used to define the propensity scores were no longer significant predictors of group status after propensity score adjustment. eTable 1 shows the study variable means after propensity score adjustment to check for the adequacy of the propensity scores to correct for baseline differences among groups. Further analyses testing for group differences within propensity score quintiles were performed, confirming the quantitative propensity score-adjusted results. The propensity score-adjusted results of Table 2 can be compared with those in eTable 2, which gives the medication- and covariate-adjusted (rather than propensity score-adjusted) results. The covariates used in eTable 2 were the same six variables included in the propensity score plus the baseline level of the change variable being analyzed. All p-values and 95% confidence intervals in the eTables are adjusted for 18 statistical tests, as was done in the main paper, by the Sidek adjustment.

Because participants who did not return for a full exam had varying completeness of medical records, we reran the results from the medication- and covariate-adjusted model (eTable 2) only on those participants attending either the Cardiovascular Genetics Clinic or the CCTS (eTable 3). This analysis excludes data obtained from physician records or self-report and uses only the highest quality data obtained at the University of Utah. We then present a multiple imputation method to infer the missing values of the lost-to-follow-up participants and any variables missing on examined participants (eTable 4).

Finally, to provide the most conservative estimates of the effects of gastric bypass surgery, we performed an intent-to-treat analysis using the post-surgical exam measurements of the controls that went on to have some form of bariatric surgery. Missing values for any of the surgical or control participants were replaced by carrying the baseline observations forward to year six (eTable 5).

Multiple imputation reduces the estimate bias caused by excluding participants with missing values.[5,6] However, multiple imputation assumes that the missing values are missing at random, which may not apply to this study, since control participants examined subsequent to gastric bypass surgery and deceased and lost-to-follow-up participants may differ from those participants participating at year 6. Data were imputed separately within each of the three study groups. The expectation-maximization algorithm was used to obtain starting estimates for a Monte Carlo Markov Chain imputation (PROC MI, SAS, Inc, Cary, NC). The multivariate normal option to impute only enough data necessary to create a monotone missing pattern was selected to create five imputation replicate datasets. This was followed by a single monotone imputation of the remaining missing variables in each of the five datasets. A mixed model was applied to each of the five imputation data sets to estimate changes from examinations one to three and test for mean differences of the changes between the surgical group and each of the two control groups separately.

Adjustment for baseline sex, age, BMI, income level, education level, marital status, and the baseline values of the dependent variable was performed to help control for differences in distributions at baseline among groups. All baseline and follow-up variables analyzed in this study were included in the multiple imputation model to prevent biased model estimates. PROC MIANALYZE was used to obtain the overall estimates and significance levels of the tests from the five datasets.

eResults

As can be seen from Table 1, the randomly-ascertained controls (control group 2) differed substantially at baseline from the other two groups. They were older and less obese at baseline. Importantly, however, the control group 2 *changes* did not differ from the control group 1 changes for any study variable (significance not shown), suggesting that six-year responses did not depend on the starting baseline levels of the participants. In addition, after propensity score adjustment, none of the weight-related variables significantly differed at baseline. Only the baseline quality of life variables for both control groups and the insulin/HOMA variables for control group 2 remained significantly different from the gastric bypass surgery group at baseline. eTable 2 shows the medication- and covariate-adjusted results, with the mean changes excluding participants lost-to-follow-up. Despite the baseline differences in age and weight, the conclusions of the study remained the same when either control group was used for comparisons to the surgery group. The very low p-values presented in this table show how significant the group differences were and suggest that extremely large biases would be required to produce such significant group differences.

Comparing the covariate adjusted results of eTable 2 to the propensity score-adjusted results of Table 2 shows that the results were very similar for the two analytical methods. Gastric bypass surgery was associated with beneficial changes for all variables measured except for the SF-36 mental component score when compared with either control group. Therefore, the findings are robust to different baseline values of the two control groups and to medication and other covariate adjustment.

The propensity-adjusted change scores of Table 2 and covariate-adjusted change scores of eTable 2 were also very similar to the covariate-adjusted change scores of eTable 3, where only participants who attended the Cardiovascular Genetics Clinic or the CCTS for a standardized examination were analyzed. These results indicate that using the medical records for participants not returning to our clinics for a full exam did not affect the conclusions of the study.

eTable 4 shows the results from a multiple imputation analysis. Again the changes and significance levels corroborate the previous tables, suggesting that any bias in excluding lost-to-follow-up participants from the analysis in Table 2 of the main paper is minimal. The validity of the missing at random assumption of the multiple imputation method is difficult to assess, but the results were not affected when those who were deceased, the group most likely to violate the assumption, were excluded from the imputation.

eTable 5 shows the intent-to-treat analysis with the baseline observation carried forward. As expected, the inclusion of approximately 100 control participants who had subsequent bariatric surgery modified the magnitude of the observed changes, but the results for all variables that were significant in Table 2 of the main paper remained significant in this analysis.

We conclude from these sensitivity analyses that the results reported in this manuscript are highly significant, and any violations of assumptions of the statistical methods have minimal effects. Gastric bypass surgery is highly effective over 6 years in improving all variables reported in this study except for the mental health component score of the SF-36.

eTable 6 presents the baseline prevalence data for the three study groups. Prevalent cases were excluded from the incidence rates and were used to determine remission rates in Table 3.

The 29 causes of death across the three study groups are provided in eTable 7. Perioperative complications that occurred less than 30 days after the 418 participants' initial gastric bypass were present in approximately 3% of the participants (data not shown), slightly less than that reported by larger, multi-center trials.[7] eTable 8 details the post-surgical (from 30 days after gastric bypass surgery to six years) hospitalizations related to bariatric-specific surgical procedures or bariatric surgery-related complications for both surgical and control participants. As shown in eTable 8, the numbers of participants hospitalized were 32 (7.7%), 22 (3.4%) and 13 (12.9%) for gastric bypass surgery patients and control participants who did and did not have subsequent bariatric surgery, respectively..

eReferences

1. **Williams RR, Hunt SC, Barlow GK, et al. Health family trees: A tool for finding and helping young members of coronary and cancer prone pedigrees in Texas and Utah. *Am J Public Health.* 1988;78:1283-1286.**
2. **Hunt SC, Williams RR, Barlow GK. A comparison of positive family history definitions for defining risk of future disease. *J Chron Dis.* 1986;39:809-821.**
3. **Heath EM, Adams TD, Daines MM, Hunt SC. Bioelectric impedance and hydrostatic weighing with and without head submersion in persons who are morbidly obese. *J Am Diet Assoc.* 1998;98(8):869-875.**
4. **Owan T, Avelar E, Morley K, et al. Favorable changes in cardiac geometry and function following gastric bypass surgery: 2-year follow-up in the Utah obesity study. *Journal of the American College of Cardiology.* 2011;57(6):732-739.**
5. **Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med.* 1991;10(4):585-598.**
6. **Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res.* 1999;8(1):3-15.**
7. **Flum DR, Belle SH, King WC, et al. Perioperative safety in the longitudinal assessment of bariatric surgery. *N Engl J Med.* 2009;361(5):445-454.**
8. **Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. *J Clin Epidemiol.* 2004;57(11):1153-1160.**
9. **Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care.* 2003;41(5):582-592.**

eTable 1. Propensity Score Adjusted Baseline Results[a] by Study Group

Outcome Variables	Gastric Bypass[b]	Control Group 1	Gastric Bypass[b]	Control Group 2
Weight, kg	132.6 (128.8-136.4) [418]	131.1 (127.4-134.9) [417]	130.0 (126.3-133.6) [418]	29.1 (124.9-133.3) [321]
Waist circumference, cm	135.2 (132.6-137.8) [418]	135.4 (132.8-138.0) [417]	133.7 (131.2-136.2) [418]	133.9 (131.0-136.8) [321]
Percent body fat, %	53.0 (52.2-53.7) [416]	52.9 (52.1-53.6) [416]	52.1 (51.4-52.8) [416]	52.1 (51.2-52.9) [310]
Systolic blood pressure, mmHg	126.4 (123.6-129.2) [418]	125.5 (122.7-128.3) [417]	127.4 (124.5-130.3) [418]	127.4 (124.0-130.8) [321]
Diastolic blood pressure, mmHg	72.1 (70.4-73.7) [418]	71.8 (70.2-73.5) [417]	72.4 (70.7-74.1) [418]	71.7 (69.8-73.7) [321]
Glucose, mg/dl	101.7 (96.3-107.0) [415]	106.5 (101.2-111.8) [417]	102.7 (97.7-107.7) [415]	105.6 (99.8-111.4) [321]
Insulin, µU/ml	19.3 (16.9-21.6) [416]	17.9 (15.6-20.3) [414]	19.1 (16.8-21.5) [416]	14.3** (11.5-17.0) [321]
HOMA-IR	4.9 (4.3-5.6) [415]	4.8 (4.1-5.5) [414]	5.0 (4.3-5.6) [415]	3.7** (2.9-4.5) [321]
HbA_{1c}, %	5.8 (5.6-6.0) [416]	6.0 (5.8-6.1) [412]	5.8 (5.7-6.0) [416]	5.9 (5.7-6.1) [319]
Total cholesterol, mg/dl	188.1 (182.7-193.6) [417]	184.4 (178.9-189.8) [417]	189.0 (183.5-194.6) [417]	187.9 (181.5-194.4) [321]
LDL-C, mg/dl	108.8 (104.6-112.9) [417]	106.7 (102.6-110.9) [416]	108.5 (104.2-112.8) [417]	109.7 (104.8-114.7) [321]
HDL-C, mg/dl	46.6 (44.8-48.3) [417]	44.8 (43.1-46.5) [416]	46.9 (45.1-48.6) [417]	46.6 (44.6-48.6) [321]
VLDL-C, mg/dl	34.3 (31.1-37.5) [417]	34.9 (31.6-38.1) [416]	35.1 (31.7-38.5) [417]	32.9 (29.0-36.9) [321]
Triglycerides, mg/dl	1878.0 (170.3-203.6) [417]	191.2 (174.5-207.9) [416]	190.6 (168.5-212.8) [417]	179.6 (154.1-205.2) [321]
IWQOL-Lite total score[c]	31.3 (28.6-33.9) [411]	35.0* (32.3-37.7) [407]	32.7 (30.0-35.5) [411]	52.8*** (49.6-56.0) [317]
SF-36 physical component score[d]	31.3 (29.9-32.8) [401]	33.4* (31.9-34.9) [400]	31.3 (29.7-32.8) [401]	39.4*** (37.7-41.2) [314]
SF-36 mental component score[e]	41.2 (39.4-43.0) [401]	40.6 (38.8-42.5) [400]	42.1 (40.2-43.9) [401]	47.0*** (44.9-49.1) [314]

* P<0.05, ** p<0.01, *** p<0.001 versus the surgical group. Two-sided p-values are adjusted for multiple comparisons.

[a] Means (95% confidence intervals) [No.].

[b] Propensity scores were derived twice, once for gastric bypass surgery versus control group 1 and once for gastric bypass surgery versus control group 2; therefore, there are different means for the gastric bypass surgery group for the two group comparisons.

[c] IWQOL-Lite, impact of weight quality of life lite, (range of scores 0-100 with 100 being best and normative mean of 94.7. A meaningful individual change is considered 7.7 to 12 points depending on baseline severity).[8]

[d] SF-36, short form 36 physical component score (range of 12-69 with 69 being best); meaningful change for either scale is 5 points with a normative mean of 50).[9]

[e] SF-36, short form 36 mental component score (range of 8-73 with 73 being best); meaningful change for either scale is 5 points with a normative mean of 50).[9]

Abbreviations: HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostasis model assessment of insulin resistance; LDL-C, measured low-density lipoprotein cholesterol; VLDL-C, measured very low-density lipoprotein cholesterol.

eTable 2. Covariate-adjusted Baseline Means [No.] (95% CI[a]) and Six-Year Study Variable Changes

Study Variables	Gastric Bypass Surgery		Control Group 1: Seeking but Did Not Have Gastric Bypass Surgery		Control Group 2: Population-based Severely Obese	
	Baseline	Change	Baseline	Change	Baseline	Change
Weight, kg	145.1 [418] (142.6 to 147.7)	-36.0 [379] (-37.9 to -34.1)	139.6*** [417] (137.1 to 142.1)	1.2*** [299] (-0.8 to 3.2)	134.2*** [321] (131.5 to 136.9)	1.2*** [296] (-0.6 to 3.1)
BMI, kg/m^2	48.3 [418] (47.5 to 49.1)	-12.9 [379] (-13.5 to -12.2)	46.7** [417] (45.9 to 47.5)	0.3*** [299] (-0.4 to 1.0)	44.8*** [321] (43.9 to 45.6)	0.3*** [296] (-0.4 to 0.9)
Waist, cm	140.8 [418] (138.9 to 142.7)	-25.6 [249] (-27.6 to -23.5)	137.9* [417] (136.1 to 139.8)	3.3*** [172] (1.0 to 5.6)	134.2*** [321] (132.1 to 136.2)	2.1*** [225] (0.1 to 4.1)
Body Fat, %	50.7 [416] (50.2 to 51.2)	-6.5 [244] (-7.0 to -5.9)	49.8** [416] (49.3 to 50.4)	-1.3*** [171] (-1.9 to -0.6)	49.0*** [310] (48.5 to 49.6)	-1.0*** [209] (-1.6 to -0.4)
SBP, mmHg	135.6 [418] (133.6 to 137.6)	-4.5 [358] (-6.5 to -2.5)	135.1 [417] (133.1 to 137.1)	4.7***[288] (2.5 to 6.9)	136.6 [321] (134.4 to 138.7)	3.3*** [293] (1.2 to 5.3)
DBP, mmHg	81.7 [418] (80.2 to 83.2)	0.0 [358] (-1.5 to 1.5)	81.4 [417] (79.9 to 82.9)	5.3*** [288] (3.7 to 6.9)	80.9 [321] (79.3 to 82.5)	4.5*** [293] (3.0 to 6.0)
Glucose, mg/dl	105.2 [415] (102.4 to 108.1)	-13.9 [336] (-17.4 to -10.4)	106.1 [417] (103.4 to 109.0)	10.0*** [262] (6.1 to 14.0)	105.2 [321] (102.2 to 108.3)	6.0*** [281] (2.4 to 9.6)
Insulin, µU/ml	15.2 [416] (14.0 to 16.5)	-10.4 [256] (-11.8 to -8.9)	13.7 [414] (12.7 to 14.9)	-2.0*** [201] (-3.6 to -0.4)	11.9*** [321] (10.9 to 13.0)	-2.8*** [248] (-4.2 to -1.4)
HOMA-IR	3.8 [415] (3.5 to 4.1)	-2.7 [253] (-3.2 to -2.3)	3.5 [414] (3.2 to 3.8)	-0.3*** [201] (-0.8 to 0.1)	3.0*** [321] (2.7 to 3.3)	-0.6*** [248] (-1.0 to -0.2)
HbA1c, %	5.8 [416] (5.7 to 5.9)	-0.3 [250] (-0.4 to -0.2)	5.9 [412] (5.8 to 5.9)	0.2*** [202] (0.1 to 0.3)	5.8 [319] (5.7 to 5.9)	0.1*** [245] (0.0 to 0.2)
Total Cholesterol, mg/dl	204.4 [417] (199.8 to 209.0)	-12.7 [295] (-17.3 to -8.1)	197.0** [417] (192.4 to 201.5)	14.9*** [255] (10.0 to 19.8)	195.2** [321] (190.3 to 200.1)	12.5*** [271] (8.0 to 17.1)
Measured LDL-C, mg/dl	124.0 [417] (120.3 to 127.8)	-7.6 [291] (-11.6 to -3.6)	119.3* [416] (115.6 to 123.0)	18.6*** [251] (14.4 to 22.8)	118.5* [321] (114.5 to 122.5)	18.8*** [270] (14.9 to 22.6)

* P<0.05, ** P<0.01, *** p<0.001 versus the surgical group adjusted for age, sex, baseline BMI (except for the baseline anthropometric variables), income, education level, and marital status. Changes also adjusted for the baseline value of the outcome variable.

[a] 95% confidence intervals (CI) and p-values are adjusted for multiple comparisons.

Abbreviations: BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HOMA-IR, homeostasis model assessment of insulin resistance; HbA1c, glycated hemoglobin; LDL-C, low-density lipoprotein cholesterol.

eTable 2 (continued). Covariate-adjusted Baseline Means [No.] (95% CI[a]) and Six-Year Study Variable Changes

Study Variables	Gastric Bypass Surgery		Control Group 1: Seeking but Did Not Have Bypass Surgery		Control Group 2: Population-based Severely Obese	
	Baseline	Change	Baseline	Change	Baseline	Change
HDL-C, mg/dl	42.3 [417] (41.1 to 43.5)	9.7 [291] (8.2 to 11.3)	40.9 [416] (39.7 to 42.0)	-4.0*** [251] (-5.6 to -2.4)	42.8 [321] (41.5 to 44.0)	-3.8*** [270] (-5.3 to -2.3)
Measured VLDL-C, mg/dl	34.8 [417] (32.9 to 36.9)	-15.3 [284] (-17.8 to -12.8)	33.6 [416] (31.8 to 35.6)	-2.3*** [239] (-4.9 to 0.4)	31.1** [321] (29.3 to 33.1)	-3.7*** [262] (-6.2 to -1.3)
Triglycerides, mg/dl	188.0 [417] (178.3 to 198.3)	-56.7 [290] (-67.8 to -45.6)	181.6 [416] (172.4 to 191.4)	3.3*** [251] (-8.4 to 15.1)	168.0** [321] (158.8 to 177.8)	1.2*** [270] (-9.6 to 11.9)
IWQOL-Lite Total Score[b]	34.9 [411] (33.0 to 36.9)	42.9 [241] (40.4 to 45.3)	38.5** [407] (36.6 to 40.4)	10.3*** [168] (7.5 to 13.1)	55.3*** [317] (53.3 to 57.3)	9.0*** [226] (6.4 to 11.6)
SF-36 Physical Component[c]	31.8 [401] (30.8 to 32.9)	11.1 [230] (9.6 to 12.5)	33.9** [400] (32.9 to 34.9)	0.6*** [167] (-1.0 to 2.2)	40.0*** [314] (38.9 to 41.1)	0.8*** [219] (-0.6 to 2.2)
SF-36 Mental Component[d]	42.7 [401] (41.4 to 44.0)	3.3 [230] (1.7 to 4.9)	41.7 [400] (40.4 to 43.0)	3.3 [167] (1.5 to 5.1)	47.4***[314] (46.0 to 48.8)	3.3 [219] (1.8 to 4.9)

* P<0.05, ** P<0.01, *** p<0.001 versus the surgical group adjusted for age, sex, baseline BMI (except for the baseline anthropometric variables), income, education level, and marital status. Changes also adjusted for the baseline value of the outcome variable.

[a] 95% confidence intervals (CI) and p-values are adjusted for multiple comparisons.

[b] IWQOL-Lite, impact of weight quality of life lite, (range of scores 0-100 with 100 being best and normative mean of 94.7. A meaningful individual change is considered 7.7 to 12 points depending on baseline severity).[12]

[c] SF-36, short form 36 physical component score (range of 12-69 with 69 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

[d] SF-36, short form 36 mental component score (range of 8-73 with 73 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

Abbreviations: HDL-C, high-density lipoprotein cholesterol; VLDL-C, very low-density lipoprotein cholesterol.

eTable 3. Baseline and Six-year Changes Only Using Participants Attending the University of Utah Clinics

Study Variables	Gastric Bypass Surgery		Control Group 1: Seeking but Did Not Have Gastric Bypass Surgery		Control Group 2: Population-based Severely Obese	
	Baseline	Change	Baseline	Change	Baseline	Change
Weight, kg	144.9 [252] (140.1 to 149.6)	-36.9 [252] (-40.4 to -33.5)	137.8* [177] (132.3 to 143.3)	0.2*** [177] (-3.6 to 3.9)	132.0*** [231] (127.3 to 136.7)	0.9*** [231] (-2.3 to 4.0)
BMI, kg/m^2	48.1 [252] (46.6 to 49.6)	-13.1 [252] (-14.2 to -11.9)	46.0 [177] (44.2 to 47.8)	0.0*** [177] (-1.3 to 1.3)	44.1*** [231] (42.5 to 45.6)	0.2*** [231] (-0.9 to 1.3)
Waist, cm	140.4 [252] (136.7 to 144.1)	-25.6 [249] (-28.8 to -22.4)	136.6 [177] (132.3 to 140.9)	3.3*** [172] (-0.4 to 7.0)	133.3*** [231] (129.6 to 137.0)	2.1*** [225] (-1.0 to 5.2)
Body Fat, %	50.5 [250] (49.5 to 51.4)	-6.5 [244] (-7.3 to -5.6)	49.4 [176] (48.3 to 50.6)	-1.3*** [171] (-2.3 to -0.3)	48.7*** [231] (47.7 to 49.7)	-1.0*** [209] (-2.0 to -0.1)
SBP, mmHg	136.4 [252] (132.4 to 140.3)	-8.1 [250] (-11.8 to -4.4)	136.5 [177] (132.0 to 141.0)	1.8*** [177] (-2.4 to 6.1)	137.1 [231] (133.2 to 141.0)	1.5*** [231] (-2.1 to 5.2)
DBP, mmHg	82.4 [252] (79.4 to 85.4)	-2.4 [250] (-5.2 to 0.4)	82.2 [177] (78.8 to 85.6)	3.4*** [177] (0.3 to 6.6)	81.3 [231] (78.4 to 84.3)	3.3*** [231] (0.6 to 6.0)
Glucose, mg/dl	103.0 [249] (97.6 to 108.7)	-14.1 [247] (-19.8 to -8.3)	105.5 [177] (99.3 to 112.2)	7.1*** [177] (0.6 to 13.5)	104.2 [231] (98.8 to 109.8)	4.0*** [231] (-1.6 to 9.5)
Insulin, μU/ml	16.0 [250] (13.5 to 18.8)	-10.3 [244] (-12.7 to -7.9)	14.4 [176] (12.0 to 17.4)	-1.1*** [170] (-3.8 to 1.6)	12.3** [231] (10.5 to 14.5)	-2.0*** [226] (-4.3 to 0.3)
HOMA-IR	4.1 [249] (3.4 to 4.9)	-2.8 [241] (-3.5 to -2.1)	3.8 [176] (3.1 to 4.7)	0.0*** [170] (-0.8 to 0.8)	3.2* [231] (2.7 to 3.8)	-0.2*** [226] (-0.9 to 0.5)
HbA1c, %	5.8 [250] (5.7 to 6.0)	-0.3 [226] (-0.5 to -0.2)	5.9 [176] (5.7 to 6.1)	0.1*** [159] (-0.1 to 0.3)	5.8 [230] (5.6 to 6.0)	0.1*** [217] (-0.1 to 0.3)
Total Cholesterol, mg/dl	203.5 [251] (194.6 to 212.4)	-13.3 [251] (-21.1 to -5.5)	192.2 [177] (182.0 to 202.3)	15.2*** [177] (6.3 to 24.1)	195.5 [231] (186.8 to 204.1)	13.3*** [230] (5.7 to 20.9)
Measured LDL-C, mg/dl	124.0 [251] (116.4 to 131.6)	-7.5 [247] (-14.1 to -0.8)	116.4 [176] (107.7 to 125.1)	20.0*** [174] (12.5 to 27.5)	118.2 [231] (110.8 to 125.6)	20.7*** [229] (14.3 to 27.1)

* P<0.05, ** P<0.01, *** p<0.001 versus the surgical group adjusted for age, sex, baseline BMI (except for the baseline anthropometric variables), income, education level, and marital status. Changes also adjusted for the baseline value of the outcome variable.

[a] 95% confidence intervals (CI) and p-values are adjusted for multiple comparisons.

[b] IWQOL-Lite, impact of weight quality of life lite, (range of scores 0-100 with 100 being best and normative mean of 94.7. A meaningful individual change is considered 7.7 to 12 points depending on baseline severity).[12]

[c] SF-36, short form 36 physical component score (range of 12-69 with 69 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

[d] SF-36, short form 36 mental component score (range of 8-73 with 73 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

Abbreviations: BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HOMA-IR, homeostasis model assessment of insulin resistance; HbA1c, glycated hemoglobin; LDL-C, low-density lipoprotein cholesterol.

eTable 3 (continued). Baseline and Six-year Changes Only Using Participants Attending the University of Utah Clinics

Study Variables	Gastric Bypass Surgery		Control Group 1: Seeking but Did Not Have Gastric Bypass Surgery		Control Group 2: Population-based Severely Obese	
	Baseline	Change	Baseline	Change	Baseline	Change
HDL-C, mg/dl	42.0 [251] (39.5 to 44.4)	10.8 [247] (8.1 to 13.4)	40.9 [176] (38.1 to 43.6)	-3.3*** [174] (-6.4 to -0.3)	43.1 [231] (40.7 to 45.4)	-3.9*** [229] (-6.5 to -1.4)
Measured VLDL-C, mg/dl	34.6 [251] (30.8 to 38.8)	-16.5 [247] (-20.4 to -12.5)	32.5 [176] (28.5 to 37.1)	-4.0*** [174] (-8.4 to 0.5)	31.2 [231] (27.9,34.9)	-4.6*** [229] (-8.4 to -0.8)
Triglycerides, mg/dl	185.7 [251] (166.3 to 207.3)	-59.9 [247] (-78.5 to -41.3)	177.5 [176] (156.6 to 201.2)	-0.3*** [174] (-21.2 to 20.6)	169.5 [231] (152.2 to 188.7)	-0.3*** [229] (-18.1 to 17.5)
IWQOL-Lite Total Score[b]	37.2 [246] (33.4 to 41.0)	43.0 [233] (39.0 to 47.0)	40.7 [172] (36.4 to 45.1)	10.2*** [159] (5.6 to 14.7)	56.8*** [229] (53.1 to 60.5)	9.4*** [220] (5.1 to 13.6)
SF-36 Physical Component[c]	32.9 [241] (30.7 to 35.0)	11.0 [222] (8.7 to 13.4)	34.7 [169] (32.3 to 37.1)	0.4*** [158] (-2.2 to 3.1)	40.1*** [225] (38.1 to 42.1)	0.9*** [212] (-1.4 to 3.2)
SF-36 Mental Component[d]	44.2 [241] (41.5 to 46.9)	3.3 [222] (0.8 to 5.9)	43.1 [169] (40.1 to 46.1)	3.1 [158] (0.2 to 6.0)	48.2** [225] (45.7 to 50.8)	3.3 [212] (0.8 to 5.7)

* P<0.05, ** P<0.01, *** p<0.001 versus the surgical group adjusted for age, sex, baseline BMI (except for the baseline anthropometric variables), income, education level, and marital status. Changes also adjusted for the baseline value of the outcome variable.

[a] 95% confidence intervals (CI) and p-values are adjusted for multiple comparisons.

[b] IWQOL-Lite, impact of weight quality of life lite, (range of scores 0-100 with 100 being best and normative mean of 94.7. A meaningful individual change is considered 7.7 to 12 points depending on baseline severity).[12]

[c] SF-36, short form 36 physical component score (range of 12-69 with 69 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

[d] SF-36, short form 36 mental component score (range of 8-73 with 73 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

Abbreviations: HDL-C, high-density lipoprotein cholesterol; VLDL-C, very low-density lipoprotein cholesterol.

eTable 4. Six-Year Changes of Study Variables by Study Group with Missing Values Assigned by Multiple Imputation

	Gastric Bypass		Control Group 1			Control Group 2		
Study Variable	**Change**	**CI**	**Change**	**CI**	**P-Value[a]**	**Change**	**CI**	**P-Value[a]**
Weight, kg	-36.3	(-39.2 to -33.4)	1.3	(-1.7 to 4.3)	5.1E-79	1.5	(-1.6 to 4.5)	5.1E-79
BMI, kg/m^2	-12.9	(-13.9 to -12.0)	0.3	(-0.7 to 1.3)	5.7E-82	0.4	(-0.7 to 1.5)	5.7E-82
Waist circumference, cm	-25.4	(-29.0 to -21.8)	4.9	(1.62 to 8.2)	9.9E-12	3.3	(0.3 to 6.4)	9.9E-12
Percent body fat, %	-6.5	(-7.6 to -5.4)	-1.4	(-3.4 to 0.6)	3.7E-16	-1.3	(-2.5 to -0.1)	3.7E-16
Systolic blood pressure, mmHg	-4.0	(-7.6 to -0.4)	5.7	(2.7 to 8.7)	5.8E-10	3.2	(-0.1 to 6.6)	1.6E-04
Diastolic blood pressure, mmHg	0.6	(-2.3 to 3.6)	6.2	(3.8 to 8.6)	1.4E-05	4.6	(2.2 to 7.1)	2.7E-02
Glucose, mg/dl	-12.8	(-18.2 to -7.4)	13.4	(3.1 to 23.7)	2.4E-08	6.7	(0.9 to 12.5)	4.1E-14
Insulin, µU/ml	-9.4	(-12.5 to -6.3)	0.8	(-5.3 to 7.0)	7.3E-03	-1.1	(-4.5 to 2.4)	1.6E-08
HOMA-IR	-2.4	(-3.9 to -1.0)	1.6	(-2.8 to 6.0)	7.3E-02	0.4	(-1.3 to 2.0)	3.0E-05
HbA_{1c}, %	-0.2	(-0.4 to -0.1)	0.3	(0.2 to 0.5)	1.4E-06	0.2	(0.0 to 0.3)	9.4E-10
Total cholesterol, mg/dl	-12.3	(-22.8 to -1.8)	15.9	(6.3 to 25.4)	2.7E-05	12.0	(4.3 to 19.7)	1.7E-07
Measured LDL-C, mg/dl	-7.4	(-18.6 to -3.9)	18.6	(9.1 to 28.2)	2.7E-04	18.0	(11.0 to 25.0)	2.6E-09
HDL-C, mg/dl	9.0	(6.8 to 11.3)	-3.9	(-7.2 to -0.5)	2.0E-11	-3.7	(-6.2 to -1.1)	2.0E-11
Triglycerides, mg/dl	-48.8	(-72.5 to -25.2)	19.8	(-16.8 to 56.3)	1.8E-05	5.4	(-21.3 to 32.1)	4.4E-06
Measured VLDL-C, mg/dl	-11.5	(-23.1 to -0.2)	2.2	(-10.1 to 14.5)	2.2E-02	-2.5	(-12.0 to 7.0)	0.34
IWQOL-Lite total score[b]	42.9	(36.4 to 49.4)	10.7	(6.4 to 15.0)	1.3E-10	9.2	(4.3 to 14.1)	1.3E-10
SF-36 physical component score[c]	2.6	(-0.6 to 5.8)	2.8	(0.6 to 5.0)	1.2E-08	3.7	(-1.3 to 3.7)	1.2E-08
SF-36 mental component score[d]	11.6	(9.6 to 13.5)	0.1	(-3.3 to 3.6)	0.99	1.2	(-0.5 to 7.9)	0.99

[a] Significance of each control group versus the surgery group. P-values and 95% confidence intervals (CI) adjusted for multiple comparisons. Sample sizes for the gastric bypass and two control groups were 418, 417, and 321 respectively. Variables measured only on the subset of participants attending the Clinical Research Center were not imputed and are not included in this table.

[b] IWQOL-Lite, impact of weight quality of life lite, (range of scores 0-100 with 100 being best and normative mean of 94.7. A meaningful individual change is considered 7.7 to 12 points depending on baseline severity).[12]

[c] SF-36, short form 36 physical component score (range of 12-69 with 69 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

[d] SF-36, short form 36 mental component score (range of 8-73 with 73 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

Abbreviations: BMI, body mass index; HOMA-IR, homeostasis model assessment of insulin resistance; HbA1c, glycated hemoglobin; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; VLDL-C, very low-density lipoprotein cholesterol.

eTable 5. Intent-to-treat Analysis of Controls who were Examined Subsequent to Gastric Bypass Surgery, with Baseline Observations Carried Forward for all Missing Values of all Participants

	Gastric Bypass		Control Group 1			Control Group 2		
Study Variable	**Change**	**CI**	**Change**	**CI**	**P-Value[a]**	**Change**	**CI**	**P-Value[a]**
Weight, kg	-32.0	(-35.4 to -28.6)	-5.6	(-8.9 to -2.3)	4.4E-78	-0.7	(-4.2 to 2.8)	1.7E-86
BMI, kg/m^2	-11.4	(-12.6 to -10.3)	-2.2	(-3.3 to -1.0)	9.1E-78	-0.5	(-1.7 to 0.7)	6.1E-85
Waist circumference, cm	-15.6	(-18.2 to -12.9)	-1.6	(-4.2 to 1.0)	1.4E-36	0.8	(-2.0 to 3.7)	2.2E-40
Percent body fat, %	-4.0	(-4.7 to -3.4)	-1.7	(-2.3 to -1.0)	2.1E-19	-0.9	(-1.6 to -0.2)	5.1E-27
Systolic blood pressure, mmHg	-4.7	(-7.6 to -1.8)	2.1	(-0.7 to 4.9)	5.7E-08	2.3	(-0.8 to 5.4)	8.2E-07
Diastolic blood pressure, mmHg	-0.8	(-3.0 to 1.3)	3.7	(1.6 to 5.8)	2.4E-06	3.8	(1.5 to 6.1)	2.5E-05
Glucose, mg/dl	-11.7	(-16.6 to -6.8)	4.9	(0.1 to 9.8)	7.9E-16	4.0	(-1.3 to 9.3)	1.9E-11
Insulin, µU/ml	-6.1	(-8.0 to -4.2)	-2.2	(-4.1 to -0.3)	7.3E-06	-2.0	(-4.1 to 0.0)	5.0E-05
HOMA-IR	-1.7	(-2.2 to -1.1)	-0.4	(-0.9 to 0.2)	5.1E-08	-0.5	(-1.1 to 0.1)	1.5E-05
HbA_{1c}, %	-0.2	(-0.3 to -0.1)	0.1	(0.0 to 0.2)	8.8E-09	0.1	(-0.1 to 0.2)	1.5E-04
Total cholesterol, mg/dl	-11.7	(-17.6 to -5.7)	6.0	(0.2 to 11.9)	2.4E-12	10.0	(3.6 to 16.5)	1.0E-14
Measured LDL-C, mg/dl	-7.6	(-12.7 to -2.5)	9.1	(4.1 to 14.1)	5.1E-15	15.3	(9.8 to 20.8)	5.5E-22
HDL-C, mg/dl	6.8	(4.8 to 8.8)	-0.9	(-2.9 to 1.1)	2.3E-20	-2.7	(-4.9 to -0.5)	1.1E-24
Triglycerides, mg/dl	-39.8	(-53.5 to -26.1)	-7.9	(-21.3 to 5.5)	1.1E-07	-0.9	(-15.7 to 13.8)	5.2E-09
Measured VLDL-C, mg/dl	-10.8	(-13.9 to -7.7)	-4.2	(-7.2 to -1.2)	1.2E-06	-3.4	(-6.7 to -0.1)	1.5E-06
IWQOL-Lite total score[b]	24.3	(20.4 to 28.1)	9.3	(5.5 to 13.0)	1.4E-20	7.7	(3.1 to 12.2)	8.3E-17
SF-36 physical component score[c]	6.0	(4.4 to 7.6)	1.6	(0.0 to 3.2)	1.7E-10	1.1	(-0.7 to 2.9)	2.5E-09
SF-36 mental component score[d]	1.8	(0.2 to 3.4)	2.2	(0.6 to 3.7)	0.99	2.8	(1.1 to 4.5)	0.99

[a] Significance of six-year change in each control group versus the surgery group. P-values and 95% confidence intervals (CI) adjusted for multiple comparisons. Sample sizes for the gastric bypass and two control groups were 418, 417, and 321 respectively. Variables measured only on the subset of participants attending the Clinical Research Center are not included in this table.

[b] IWQOL-Lite, impact of weight quality of life lite, (range of scores 0-100 with 100 being best and normative mean of 94.7. A meaningful individual change is considered 7.7 to 12 points depending on baseline severity).[12]

[c] SF-36, short form 36 physical component score (range of 12-69 with 69 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

[d] SF-36, short form 36 mental component score (range of 8-73 with 73 being best); meaningful change for either scale is 5 points with a normative mean of 50).[13]

Abbreviations: BMI, body mass index; HOMA-IR, homeostasis model assessment of insulin resistance; HbA1c, glycated hemoglobin; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; VLDL-C, very low-density lipoprotein cholesterol.

eTable 6. Prevalence Rates of Diabetes, Hypertension, and Dyslipidemia by Study Group

	Prevalence, % (95% CI)[a] [No./Total]		
Clinical Endpoint	**Gastric Bypass**	**Control Group 1**	**Control Group 2**
Diabetes	22 (17-27%) [93/418]	25 (20-30%) [106/417]	29 (22-36%) [92/321]
Hypertension	43 (37-49%) [181/418]	43 (37-49%) [179/417]	52 (45-59%) [167/321]
Low HDL-C	41 (35-47%) [172/418]	45 (39-51%) [187/417]	36 (29-43%) [114/321]
High LDL-C	20 (15-25%) [84/418]	18 (13-23%) [75/417]	21 (15-27%) [66/321]
High Triglycerides	43 (37-49%) [181/418]	41 (35-47%) [172/417]	41 (34-48%) [130/321]

[a] CI: confidence interval, adjusted for five multiple comparisons.

eTable 7. Causes of Death by Study Group

Causes of Death	Gastric Bypass Surgery Patients	Control Group 1: Seeking but Did Not Have Gastric Bypass Surgery	Control Group 2: Population-based Severely Obese
Aneurysm	0	1	0
Coronary Artery Disease	0	1	0
Cancer	2	2	1
Chronic Obstructive Pulmonary Disease	0	1	0
Diabetes	1	0	0
Heart Failure	0	2	1
Hypertension	0	1	0
Pulmonary Hypertension	0	1	0
Obesity[a]	0	1	0
Pulmonary Embolism	0	0	1
Renal Failure	0	1	0
Stroke	0	1	0
Poisoning of Undetermined Intent	2	1	0
Suicide	4	0	0
Other	3	1	0
Total	**12**	**14**	**3**

[a] ICD code E66.8 for "other: obesity." No secondary cause of death listed.

eTable 8. Hospitalizations with bariatric surgery-related ICD-9 codes

	Number of Hospitalizations (% of group total)		
		Combined Control Groups	
	Surgery Group	No Weight Loss Surgery	Subsequent Weight Loss Surgery
Complications	[n=418]	[n=638]	[n=101]
Intestinal adhesion/obstruction	14 (3.3)	5 (0.8)	4 (4.0)
Acute or chronic cholecystitis	9 (2.1)	9 (1.4)	2 (2.0)
Hernias	4 (1.0)	5 (0.8)	2 (2.0)
Ulcers	2 (0.5)	1 (0.2)	9 (8.9)
Acute vascular insufficiency of intestine	2 (0.5)	---	---
Pulmonary embolism	---	3 (0.5)	1 (1.0)
Other complications of procedures not classified elsewhere	7 (1.7)	3 (0.5)	7 (6.9)
Total hospitalizations[a]	38 (9.1)	26 (4.1)	25 (24.8)
Total participants hospitalized	32 (7.7)	22 (3.4)	13 (12.9)

[a] Hospitalizations after 1st exam or >30 days after gastric bypass surgery to 6-year follow-up