Supplementary Appendix

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

Supplement to: Adams TD, Davidson LE, Litwin SE, et al. Weight and metabolic outcomes 12 years after gastric bypass. N Engl J Med 2017;377:1143-55. DOI: 10.1056/NEJMoa1700459

Weight and Metabolic Outcomes 12 Years Following Gastric Bypass Surgery

Supplementary Appendix

Table of Contents **Page Number**

List of Investigators

Ted D. Adams, PhD, MPH. Intermountain LiVe Well Center Salt Lake, Intermountain Healthcare and Division of Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.
Lance E. Davidson, PhD. Department of Exercise Sciences, Brigham Young University, Provo, Utah and Division of Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.
Sheldon E. Litwin, MD. Medical University of South Carolina, and Ralph H. Johnson VA Medical Center, Charleston, South Carolina.
Jaewhan Kim, PhD. Department of Health, Kinesiology and Recreation, College of Health, University of Utah, Salt Lake City, Utah.
Ronette L. Kolotkin, PhD. Quality of Life Consulting, Durham, North Carolina; Department of Community and Family Medicine, Duke University Health System, Durham, North Carolina; Sogn og Fjordane University College, Department of Health Studies, Førde, Norway; Department of Surgery, Førde Central Hospital, Norway; Førde Hospital Trust, Førde, Norway; and Morbid Obesity Centre, Vestfold Hospital Trust, Tønsberg, Norway.
M.N. Nanjee, PhD. Division of Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.
Jonathan M. Gutierrez, BS. Division of Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.
Sara J. Frogley, MBA. Department of Genetic Medicine, Weill Cornell Medicine, Doha, Qatar and Division of Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.
Anna R. Ibele, MD. Division of General Surgery, Department of Surgery, University of Utah School of Medicine, Salt Lake City, Utah.
Paul N. Hopkins, MD, MSPH. Divisions of Cardiovascular Medicine and Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.
Eliot A. Brinton, MD. Utah Foundation for Biomedical Research and Utah Lipid Center, Salt Lake City, Utah.
Rodrick McKinlay, MD. Rocky Mountain Associated Physicians, Salt Lake City, Utah.
Steven C. Simper, MD. Rocky Mountain Associated Physicians, Salt Lake City, Utah.
Steven C. Hunt, PhD. Department of Genetic Medicine, Weill Cornell Medicine, Doha, Qatar and Division of Cardiovascular Genetics, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah.

Methods

<u>Selection and rationale for two non-surgical groups</u>

In addition to the gastric bypass surgery group, we designed the study to have two non-surgery comparison groups, each with different selection criteria.[1] The purposes of the two groups were to be able to make inferences to a larger population of all subjects seeking gastric bypass surgery and to a sample of people who were severely obese, but were not necessarily seeking surgery.

We began the study by recruiting all patients who were actively seeking gastric bypass surgery from the participating surgical practice. Once enrolled, patients underwent a baseline examination and later, when the patient learned whether or not their insurance company would cover the surgical procedure, they were assigned to the surgical or non-surgical groups. There were very few patients who chose not to have the surgery after their insurance approved the procedure, or who chose to “self-pay” after their insurance denied coverage. For almost all of the 835 patients seeking gastric bypass, the deciding factor as to whether or not they had surgery was based on whether the specific insurance plan offered by the patient’s employer covered bariatric surgery. When this study began, insurance company plans did not cover bariatric surgery to the extent that they do today. Therefore, while there were no matching criteria per se, other than having a BMI ≥ 35 kg/m^2 and qualifying for gastric bypass surgery, the subjects were implicitly matched on characteristics that led them to seek weight-loss surgery. Non-surgery group 1 is the most appropriate comparison group to use for inferring how effective gastric bypass surgery is for those who seek surgery.

The second comparison group was a random sample of the Utah population who were severely obese and selected only for a BMI ≥ 35 kg/m^2. This group (non-surgery group 2) allowed us to infer how effective gastric bypass surgery is compared with the more general population who is severely obese. Because this group was not matched, it is more prone to unmeasured confounding. Use of two groups with different characteristics and selection criteria also provided an indication of the likelihood that unmeasured confounding that could not be controlled by covariate adjustment affected our results and conclusions.

Statistical Methods

In addition to the primary endpoints mentioned in the main paper, secondary endpoints included risk factors for the primary endpoints or variables used in the definition of the primary endpoints. The secondary endpoint results are listed in Tables 2 and S2. The primary endpoint results are shown in Table 3 and Figure 2.

Table 1 shows the unadjusted means, 95% confidence intervals, and sample sizes for the baseline and 12-year exams for the main variables of the study. Statistics for the non-surgical groups are shown with and without the participants who had bariatric surgery subsequent to their baseline exam. No adjustments for covariates or for multiple comparisons were used in this table. Table S1 shows the same data for additional variables collected in this study.

General Linear Model Adjustment

Adjustments of the variables listed in Tables 2 and 3 for possible confounding variables were performed using a general linear model with study group as the independent variable and six covariates. For the previously published six-year follow-up analysis, a propensity score was derived from an initial large number of variables collected at baseline.[2] There were six variables that predicted differences among the three study groups at baseline and were retained in the propensity score: sex, age, baseline BMI, marital status (married, not currently married), income, and years of education. Because there were only a few variables significantly predicting group status, the propensity score-adjusted results were nearly identical to the covariate-adjusted results. Therefore, the simpler covariate adjustment method was used for this study. The six variables defining the propensity score were included in each model as individual covariates.

To help assess adequacy of the covariate adjustments to remove differences among groups, we tested for baseline differences in multiple other variables collected, but not included, in the 12-year analysis of this study, after adjustment for the 6 covariates. Because of the broad representation of multiple physiological systems covered by these variables, they also help indicate the likelihood that unmeasured confounding was a significant factor related to this study (Table S3). This likelihood is discussed further in the following discussion section.

Multiple Comparison Adjustment

There are 15 variables listed in Tables 2 and S2 that include two non-surgery group comparisons versus the surgery group. While a conservative adjustment for multiple comparisons would be based on 30 statistical tests, many of the variables are highly correlated. The variables represent 5 general categories (weight, blood pressure, lipids, glucose/insulin, and quality of life), resulting in about 10 statistical tests. Because the traits are only partially correlated on the one hand, and to prevent overcorrection for 30 tests on the other hand, a compromise was made by assuming that the analysis represents approximately 15 independent tests, the actual number of variables. The required alpha to reach a nominal statistical significance of 0.05 for 15 tests is 0.0034. The confidence intervals were adjusted using the Sidak method based on an alpha level of 0.0034 and the adjusted intervals are presented in the table.[3] The Sidak method is not quite as conservative as the Bonferroni correction, but the differences are minimal (alpha level of 0.0034 versus 0.0033). The Sidak method was used to be comparable to our methodology published for the 6-year results.

Table 3 shows the primary incidence and remission variables for five somewhat correlated endpoints. For this set of analyses, we assumed 10 independent tests and set the alpha level to calculate the 95% confidence intervals at 0.0051, again based on the Sidak formula.[3]

SAS version 9.4 (SAS Institute, Cary, NC) was used for all analyses.

Results

Evaluation of Bias and Confounding

Table 1 shows that the surgery and non-surgery groups were indeed very similar for baseline characteristics of the primary variables. Table S1 shows similar data for other variables included in this study. We had previously shown that there were no significant differences in baseline prevalence of diabetes, hypertension, stroke, coronary events or cancer among the groups.[1]

At baseline, the non-surgery group 2 participants were older, weighed less, had smaller waist circumferences, and reported higher quality of life than the surgery group. Lipid levels and blood pressures were not different, while fasting glucose and HbA1c were slightly higher (Tables 1 and S1).

Table S3 shows the comparison of baseline means among groups after adjustment for the 6 covariates included in all statistical models. The measurement and definitions of these variables have been published.[1] The p-values are unadjusted for multiple comparisons. Only HDL-C, insulin, and HbA_{1c} showed evidence of significant mean differences (<3%, <5% and <7%, respectively) between the surgical and the non-surgical group 1. Fitness, as assessed by a treadmill test, and calorie intake, as assessed by the Willett semi-quantitative questionnaire, were not different. No difference was seen for sleep apnea, pulmonary function, or echocardiographic measurements.

Health-related quality of life

Twelve-year changes in quality of life assessed by the IWQOL-Lite total score and SF-36 physical component summary score were significantly improved (p<0.001) in the surgery group compared with non-surgery groups 1 and 2 (Table S2). However, there were no significant 12-year group differences for the SF-36 mental component summary score between the surgery group and non-surgery group 1 (p=0.40 excluding and p=0.99 including subsequent bariatric surgery subjects) or group 2 (p=0.37

excluding and p=0.39 including subsequent bariatric surgery subjects). Table S2 shows similar 12-year change data for other variables included in this study.

Mortality

Mortality at 12-years follow-up was 6.2% (26 deaths), 9.4% (39 deaths) and 5.6% (18 deaths) for the RYGB group and non-surgery groups 1 and 2, respectively. After excluding the non-surgery group subjects who later went on to have bariatric surgery, age- and sex-adjusted mortality for the RYGB group was significantly lower compared with the non-surgery group 1 (OR=0.53 [95% CI, 0.30, 0.92]; p=0.02), but not significantly different than the non-surgery group 2 (OR=1.32 [95% CI; 0.70, 2.51]; p=0.39). There were 16 cancer and 28 cardiovascular-related deaths in all groups combined.

The only pre-specified hypothesis regarding mortality was the total mortality would be reduced after gastric bypass surgery. It was not anticipated that the number of cause-specific deaths would lead to sufficient statistical power for further testing. However, due to suggestions by reviewers, post-hoc testing of the association of study group with suicide and poisoning deaths was performed. Fisher's exact test was used for all tests and no adjustment for multiple comparisons was performed. Table S4 shows the number of deaths for each group.

Table S5 shows the statistical tests for various combinations of causes of death and study groups. When the two non-surgical groups were combined, both the association with suicide deaths (p=0.015) and suicide plus poisoning deaths (p=0.013) were significant. However, the significant effects mostly resulted from the lack of deaths in the non-surgery group 2.

Discussion

Evaluation of Bias and Confounding

Because this study was an observational study and not a randomized trial, either bias from comparison group selection or unmeasured confounding even after covariate adjustment could have influenced the results and conclusions.

An important reason for studying the two different non-surgical groups was to help assess bias induced by the selection criteria. The non-surgery group who was seeking gastric bypass surgery appeared very representative of those who went on to have surgery. The second non-surgical group with severe obesity was intentionally selected to represent the more general population. Despite different baseline characteristics, including BMI, of the two non-surgical groups, the study results at 2, 6 and 12 years consistently demonstrated that gastric bypass surgery is effective for weight loss and improvement of diabetes, hypertension, and dyslipidemia. Further, the small differences in effect sizes between the two non-surgical group comparisons to the surgical group suggested that whatever unmeasured biases or unadjusted confounding existed in this study, the effects appear to have been minimal or at least robust to that confounding. The conclusions of the study remained the same regardless of which non-surgical group was used to compare to the surgical group. The magnitude of the effect sizes suggests that unmeasured or unadjusted biases would have to be very large to make the estimates non-significant.

Our prior study of the 6-year exam used propensity scores to adjust for baseline differences. In that paper, we showed that after selecting the variables that predicted baseline group membership and using them as covariates in the analytical models resulted in similar results and conclusions.[2] Since two propensity scores were derived in the prior paper, one for each non-surgical group, this method was not as efficient as covariate adjustment. Therefore, covariate adjustment for the six baseline variables was used in the current analysis. While unmeasured confounding could still exist, baseline analysis of the large number of variables measured in this study did not suggest evidence for this confounding and it is unclear what variables that were not measured would have a large enough impact on the group comparisons to alter the conclusions. Table S3 in particular shows that when we expand the variables analyzed that were available in a subset of the participants with specialized testing, few of these variables showed baseline differences. Even though HbA_{1c} was significantly higher in the non-surgery group 1 than in the surgery group, it was less than 3% higher, not enough to explain our diabetes findings. The lack of a baseline difference in the non-surgery group 2 with similar diabetes remission and incidence at 12 years to that in the non-surgery group 2, also suggests lack of confounding due to this variable. In addition, fasting insulin levels were <5% lower, suggesting a little less insulin resistance despite the slightly higher HbA_{1c}. Since the expanded variable list covers the aspects of the heart, sleep apnea, fitness, calorie intake, metabolic rate, and pulmonary function, lack of baseline differences in these variables diminishes the likelihood that unmeasured confounding could explain the very large effect sizes seen after surgery, especially when using two different non-surgical groups as comparison groups. Randomized studies would need to be performed to ultimately prove that our findings are not confounded in some way.

The results of this investigation apply to the mostly non-Hispanic white population used in this study. The diabetes findings might not represent those found by studying Hispanics, for example, or the blood pressure findings may not apply to the African-American population, who suffer from a greater burden of hypertension.

<u>Mortality</u>

The number of suicides, while small numerically, are concerning as a percentage of the study size. However, statistical power is low for testing cause-specific mortality. The tests in Table S4 are all post-hoc tests and were they adjusted for multiple comparisons would not be significant. However, the direction of the effects from this prospective study are similar to the significant effects seen in our prior retrospective study of mortality where it was shown that suicide, poisonings and accidental deaths were significantly related to gastric bypass surgery. Since publication of our original mortality study, other large studies have confirmed the findings with regards to self-harm.[4-7] An important question is whether or not bariatric surgery is causal for suicide or just coincident with it. We note that at baseline, the surgical group had significantly worse quality of life than either nonsurgical group. This has been seen in other studies,[8] suggesting that it may not be the surgery per se that leads to suicide, but the quality of life before surgery that was not sufficiently improved by the surgery from the patients' standpoint that might lead to suicide. It is possible that the combination of prior psychological problems combined with over-expectations that they would improve after surgery may lead to these deaths. Findings to date suggest greater attention needs to be given to determine the underlying symptoms and predictors of suicide in those seeking bariatric surgery.

The total mortality rate for the RYGB group was 2/3 that of the non-surgical group 1, in line with our larger mortality study.[7] The non-surgery group 2 had a similar mortality as the RYGB group, likely resulting from their being healthier at baseline. Cause of death from the NDI records is not 100% accurate and differential assignment of cause of death on the death certificates due to the presence or absence (after surgery) of severe obesity cannot be ruled out. However, the numbers of deaths in this study is small enough that firm conclusions about mortality should be made with extreme caution and should rely on the larger studies published.

Table S1. Baseline and 12-year results for variables in this study not included in Table 1.

Study Variables	RYGB Surgery Group		Non-surgery Group 1			Non-surgery Group 2		
	Baseline exam	12 Year Exam	Baseline Exam	12 Year Exam		Baseline Exam	12 Year Exam	
				Excluding Subsequent Bariatric Surgery Patients	Including Subsequent Bariatric Surgery Patients		Excluding Subsequent Bariatric Surgery Patients	Including Subsequent Bariatric Surgery Patients
	Mean (95% CI) No.	Mean (95% CI) No.	Mean (95% CI) No.	Mean (95% CI) No.	Mean (95% CI) No.	Mean (95% CI) No.	Mean (95% CI) No.	Mean (95% CI) No.
Follow-up time, years	--	11.8 (11.7, 11.9) 388	--	11.7 (11.6, 11.9) 217	11.7 (11.6, 11.8) 364	--	11.8 (11.7, 11.9) 262	11.8 (11.7, 11.9) 301
White non-Hispanics, N	386/418	--	388/417	--	--	319/321	--	--
BMI, kg/m^2	47.2 (46.5, 47.9) 418	35.2 (34.3, 36.1) 387	46.2 (45.5, 46.9) 417	45.0*** (43.8, 46.1) 217	41.8*** (40.9, 42.7) 363	43.8*** (43.0, 44.6) 321	44.0*** (42.9, 45.1) 262	43.2*** (42.1, 44.2) 301
Waist circumference, cm	136.0 (134.3, 137.6) 418	113.7 (111.1, 116.3) 254	134.5 (132.8, 136.1) 417	134.3*** (130.8, 137.9) 137	127.2*** (124.4, 130.0) 235	130.8*** (129.0, 132.7) 321	133.9*** (130.9, 136.8) 199	131.5*** (128.6, 134.3) 229
Insulin, µU/ml	19.3 (17.9, 20.8) 416	9.1 (6.4, 11.7) 252	17.9 (16.4, 19.3) 414	17.2*** (13.5, 20.9) 135	13.7* (11.2, 16.3) 233	14.0*** (12.3, 15.6) 321	21.2*** (18.2, 24.2) 199	19.6*** (17.0, 22.2) 229
HOMA-IR	5.0 (4.5, 5.4) 415	2.1 (1.1, 3.0) 252	4.8 (4.4, 5.2) 414	5.2*** (3.9, 6.5) 135	3.9** (3.0, 4.8) 233	3.7*** (3.3, 4.2) 321	6.3*** (5.2, 7.3) 199	5.8*** (4.8, 6.7) 229
IWQOL-Lite, total score	31.4 (29.6, 33.1) 411	71.9 (69.4, 74.5) 254	34.9** (33.1, 36.6) 407	46.9*** (43.4, 50.4) 135	56.7*** (53.9, 59.5) 235	54.5*** (52.5, 56.5) 317	61.9*** (59.0, 64.8) 199	62.8*** (60.0, 65.6) 231
SF-36, physical component score	31.4 (30.4, 32.3) 401	42.3 (41.0, 43.6) 251	33.3** (32.4, 34.3) 400	36.0*** (34.2, 37.7) 137	39.2** (37.8, 40.5) 236	39.3*** (38.2, 40.3) 314	37.5*** (36.0, 38.9) 199	38.1*** (36.7, 39.5) 228
SF-36, mental component score	41.4 (40.3, 42.6) 401	47.0 (45.6, 48.4) 251	40.4 (39.3, 41.6) 400	43.6** (41.7, 45.5) 137	45.2 (43.8, 46.7) 236	47.7*** (46.5, 49.0) 314	51.4*** (49.8, 53.0)) 199	51.2*** (49.7, 52.7) 228

Abbreviations: RYGB, Roux-en-Y gastric bypass; BMI, body mass index; HOMA-IR, homeostatic model assessment of insulin resistance; IWQOL-Lite, Impact of Weight Quality of Life-Lite; SF-36, 36-Item Short Form Health Survey.
Range of scores for IWQOL-Lite: 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.[9]
Range of scores for SF-36 physical component score: 12-69, with 69 being best); meaningful change is 5 points with a normative mean of 50.[10]
Range of scores for SF-36 mental component score: 8-73, with 73 being best); meaningful change is 5 points with a normative mean of 50.[10]
No adjustment of the confidence intervals or significance levels was performed for multiple comparisons.
* P < 0.05. (Compared with the surgical group for baseline and 12 year follow-up.)
** P < 0.01. (Compared with the surgical group for baseline and 12 year follow-up.)
*** P < 0.001. (Compared with the surgical group for baseline and 12 year follow-up.)

Table S2. Adjusted 12-Year Change from Baseline by Study Group and Group Differences.†

Study Variables	Surgery Group Mean (95% CI)	Non-surgery Group 1 Mean (95% CI)		Non-surgery Group 2 Mean (95% CI)	
		Excluding Subsequent Bariatric Surgery Patients	Including Subsequent Bariatric Surgery Patients	Excluding Subsequent Bariatric Surgery Patients	Including Subsequent Bariatric Surgery Patients
Waist circumference, cm	-19.6 (-23.1, -16.2)	4.5*** (0.1, 9.0)	-5.1*** (-9.1, -1.1)	5.6*** (2.0, 9.3)	2.0*** (-1.9, 5.9)
Insulin, µU/ml	-9.4 (-14.4, -4.5)	-3.0 (-9.4, 3.4)	-6.2 (-10.8, -1.6)	0.9*** (-4.3, 6.0)	-0.8*** (-5.2, 3.7)
HOMA-IR	-2.7 (-4.5, -1.0)	-0.4 (-2.6, 1.8)	-1.6 (-3.2, 0.0)	0.9*** (-0.9, 2.7)	0.3*** (-1.2, 1.9)
IWQOL-Lite, total score[4]	35.5 (31.5, 39.5)	8.5*** (3.4, 13.6)	19.3*** (14.9, 23.7)	10.7*** (6.2, 15.3)	14.0*** (9.4, 18.7)
SF-36, physical component score	8.1 (6.0, 10.2)	1.3*** (-1.3, 3.9)	4.6** (2.5, 6.7)	0.6*** (-1.6, 2.8)	1.8*** (-0.3, 4.0)
SF-36, mental component score	3.3 (0.9, 5.7)	0.8 (-2.3, 3.9)	2.6 (0.3, 5.0)	5.7 (3.1, 8.2)	5.9 (3.6, 8.3)

Abbreviations: HbA1c, hemoglobin A_{1c}; HOMA-IR, Homeostatic Model Assessment of Insulin Resistance; IWQOL-Lite, Impact of Weight Quality of Life - Lite; SF-36, 36-Item Short Form Health Survey.
†For Non-surgery Groups 1 and 2, two columns are included; first column excludes participants who chose to later have any type of bariatric surgical procedure and second column includes all non-surgery participants.
Confidence intervals and significance levels were adjusted for multiple comparisons (see supplementary appendix).
* P < 0.05. (Compared with the surgical group for 12 year follow-up.)
** P < 0.01. (Compared with the surgical group for 12 year follow-up.)
*** P < 0.001. (Compared with the surgical group for 12 year follow-up.)

Table S3. Comparisons of adjusted baseline means and standard errors between groups.

Variable	RYGB Surgery Mean (SE)	Non-Surgery Group 1 Mean (SE)	Non-Surgery Group 2 Mean (SE)
Height, cm	173.6 ± 0.37	173.0 ± 0.38	172.9 ± 0.41
Waist Circumference, cm	136.9 ± 0.65	136.6 ± 0.64	136.0 ± 0.69
Systolic Blood Pressure, mm Hg	129.4 ± 1.10	128.8 ± 1.02	129.3 ± 1.04
Diastolic Blood Pressure, mm Hg	74.0 ± 0.63	74.0 ± 0.62	73.2 ± 0.67
LDL-C, mg/dL	106.4 ± 1.59	104.6 ± 1.57	108.3 ± 1.69
HDL-C, mg/dL	44.2 ± 0.61	42.2 ± 0.60**	43.7 ± 0.64
Triglycerides, mg/dL	196.0 ± 8.0	199.7 ± 7.9	185.0 ± 8.5
Glucose, mg/dl	105.1 ± 1.99	109.3 ± 1.96	108.1 ± 2.11*
Insulin, µU/ml	20.1 ± 0.85	18.8 ± 0.83**	15.2 ± 0.90***
Hemoglobin A_{1c}, %	5.87 ± 0.06	6.03 ± 0.06*	5.95 ± 0.06
Left Ventricular Mass, g	201.1 ± 3.23	196.5 ± 3.28	208.7 ± 3.18**
Ejection Fraction, %	0.52 ± 0.04	0.62 ± 0.04	0.58 ± 0.04
Cardiac Output, L/min	4.61 ± 0.20	4.78 ± 0.21	4.58 ± 0.20
Respiratory Disturbance Index, hour	27.0 ± 1.63	27.1 ± 1.64	27.9 ± 1.58
Treadmill Time, min	618 ± 12	595 ± 12	609 ± 12
Total Calories, Kcal	2187 ± 67	2168 ± 67	2194 ± 72
Resting Metabolic Rate, kcal per day	2336 ± 20	2391 ± 20*	2310 ± 21
Forced Expiratory Volume at 1 sec, L	3.20 ± 0.03	3.16 ± 0.03	3.15 ± 0.03

Abbreviations: RYGB, Roux-en-Y gastric bypass; SE, standard error of the mean

* P <0.05; ** P <0.01; *** P <0.001 compared with the surgical group

Table S4. Suicide and Poisoning Deaths with Gastric Bypass Surgery

	Deaths (%) in Surgery Group	Deaths (%) in Non-surgery Group 1	Deaths (%) in Non-surgery Group 2
Total Deaths	26 (6%)	39 (9%)	18 (6%)
Suicide	5 (1.2%)	2 (0.5%; both after RYGB)	0
Poisoning	4 (1.0)	2 (0.5%; both prior to RYGB)	0

Abbreviations: RYGB, Roux-en-Y gastric bypass

Table S5. Significance of Suicide and Poisoning Deaths with Gastric Bypass Surgery

Comparison	Suicide deaths	Suicide and Poisoning Deaths
Surgery vs combined non-surgical groups	P=0.015	P=0.013
Surgery vs non-surgical group 1	P=0.16	P=0.22
Surgery vs non-surgical group 2	P=0.09	P=0.013

References

1. Adams TD, Avelar E, Cloward T, et al. Design and rationale of the Utah obesity study. A study to assess morbidity following gastric bypass surgery. Contemp Clin Trials 2005;26:534-51.
2. Adams TD, Davidson LE, Litwin SE, et al. Health benefits of gastric bypass surgery after 6 years. JAMA 2012;308:1122-31.
3. Sidak M. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. J Am Stat Assoc 1967;62:626-33.
4. Peterhansel C, Petroff D, Klinitzke G, Kersting A, Wagner B. Risk of completed suicide after bariatric surgery: a systematic review. Obes Rev 2013;14:369-82.
5. Tindle HA, Omalu B, Courcoulas A, Marcus M, Hammers J, Kuller LH. Risk of suicide after long-term follow-up from bariatric surgery. Am J Med 2010;123:1036-42.
6. Bhatti JA, Nathens AB, Thiruchelvam D, Grantcharov T, Goldstein BI, Redelmeier DA. Self-harm Emergencies After Bariatric Surgery: A Population-Based Cohort Study. JAMA surgery 2016;151:226-32.
7. Adams T, Hunt S. Long-term mortality following gastric bypass surgery. Presented at the International Congress of Obesity meetings, Australia 2006.
8. Mitchell JE, Crosby R, de Zwaan M, et al. Possible risk factors for increased suicide following bariatric surgery. Obesity (Silver Spring) 2013;21:665-72.
9. Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. J Clin Epidemiol 2004;57:1153-60.
10. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable university of half a standard deviation. Med Care 2003;41:582-92.