Supplementary Appendix

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

Supplement to: Schauer PR, Bhatt DL, Kirwan JP, et al. Bariatric surgery versus intensive medical therapy for diabetes — 5-year outcomes. N Engl J Med 2017;376:641-51. DOI: 10.1056/NEJMoa1600869

Supplementary Appendix

Supplement to: Schauer PR, Bhatt DL, Kirwan JP, et al. Bariatric Surgery versus Intensive Medical Therapy for Diabetes: 5-year Outcomes. N Engl J Med 2017.

Table of Contents Page

Author Contributions:

Drs. Schauer, Kashyap, and Bhatt had full and independent access to all of the data and vouch for the integrity and the accuracy of the analysis.

Study concept and design: Schauer, Kashyap, Brethauer, Bhatt

Acquisition of data: Schauer, Kashyap, Aminian, Kim, Pothier, Wolski

Analysis and interpretation of data: Schauer, Kashyap, Wolski, Bhatt, Nissen, Aminian, Navaneethan, Kirwan, Brethauer, Singh

Drafting of the manuscript: Schauer, Kashyap, Pothier, Wolski, Aminian, Navaneethan, Sing,

Critical revision of the manuscript for important intellectual content: Nissen, Bhatt, Kirwan, Aminian, Navaneethan, Kashyap, Brethauer

Statistical analysis: Wolski

Obtained funding: Schauer, Kirwan, Kashyap

Administrative technical or material support: Schauer, Kashyap, Pothier, Bhatt

Study supervision: Schauer, Kashyap, Bhatt

Financial support: This research was funded by a grant from Ethicon, and conducted with support from the Investigator-Initiated Study Program of LifeScan Inc., Cleveland Clinic and NIH.

Role of the sponsor: The sponsor participated in discussions regarding study design and protocol development. The database, statistical analysis, and monitoring were all performed by the Cleveland Clinic Coordinating Center for Clinical Research. The manuscript was prepared by the corresponding author and modified after consultation with co-authors. The sponsor was permitted to review the manuscript and suggest changes, but the final decision on content and submission was exclusively retained by the academic authors.

Data and Safety Monitoring Board

J. Michael Henderson, MD (Chair); James B. Young, MD; Venu Menon, MD, Cleveland Clinic

Study Governance

Drs. Schauer, Kashyap, and Bhatt designed the trial with advice from Dr. Nissen. The study was approved by the Cleveland Clinic Institutional Review Board, and all subjects provided written informed consent. The Cleveland Clinic Coordinating Center for Clinical Research gathered and analyzed the data. An independent Data and Safety Monitoring Board reviewed the safety and conduct of the trial at confidential meetings held yearly. The first author vouches for the integrity of the data, wrote the first draft of the manuscript, and prepared subsequent drafts with input from all the coauthors. All the authors made the decision to submit the manuscript for publication. There were no agreements concerning confidentiality of the data between the funding sponsor and the authors or the institutions named in the credit lines.

Additional Statistical Methods

We report all continuous variables with a normal distribution as means and standard deviations. Variables with a non-normal distribution are reported as medians and interquartile ranges. Categorical variables are summarized with the use of frequencies.

A logistic model adjusting for insulin use at baseline was used to generate pairwise p-values. P-values are reported with and without adjustment for multiplicity using the stepdown Bonferroni method. P-values for the secondary endpoints were similarly adjusted for multiplicity. To assess the impact of the 16 patients missing final HbA1c information, the Multiple Imputation procedure in SAS (PROC MI) was used to impute missing data using the fully conditional specification method on the imputed population. All available hemoglobin data, demographics, medical history, laboratory data (glucose, lipids) and blood pressure were included as predictors in the imputation model. A logistical model including terms for treatment and insulin strata was run on 20 imputed datasets and results were combined using PROC MIANALYZE. The inverse-link transformation was used to obtain event probabilities. The resulting pairwise p-values were then adjusted according to the step-down Bonferroni procedure using PROC MULTTEST.

Supplementary Tables

Table S1. Characteristics of the Patients at Baseline*

Parameter	Medical Therapy (N=38)	Gastric Bypass (N=49)	Sleeve Gastrectomy (N=47)	P Value
Duration of diabetes - yrs	8.8 ± 5.5	8.2 ± 5.6	8.3 ± 4.6	0.74
Insulin Users – no. (%)	17 (44.7)	22 (44.9)	20 (42.6)	0.97
Age – yrs	50.2 ± 7.7	48.2 ±8.5	48.1 ± 8.1	0.41
Female sex – no. (%)	25 (65.8)	28 (57.1)	36 (76.6)	0.13
Body-mass index – (kg/m^2)	36.4 ± 3.0	37.0 ± 3.4	36.0 ± 3.9	0.33
Body-mass index <35	17 (44.7)	14 (29.1)	18 (38.3)	0.29
Body weight - kg	105.0 ± 14.4	106.8± 14.9	100.4 ± 16.8	0.12
Waist circumference - cm	113.7 ± 9.4	116.5 ± 9.2	113.5 ± 10.3	0.25
Waist to hip ratio	0.95 ± 0.09	0.96 ± 0.07	0.95 ± 0.09	0.74
White race – no. (%)ǂ	27 (71.1)	36 (73.5)	34 (72.3)	0.97
Smoker – no. (%)	12 (31.6)	19 (38.8)	10 (21.3)	0.17
Metabolic syndrome – no. (%)	34 (89.5)	45 (91.8)	45 (95.7)	0.55
History of dyslipidemia – no. (%)	32 (84.2)	43 (87.8)	38 (80.9)	0.65
History of hypertension – no. (%)	24 (63.2)	35 (71.4)	28 (59.6)	0.46

* Plus-minus values are means ± SD. P values are for the overall comparisons.
ǂ Race was self-reported.

Table S2. Factors Associated with the Primary Endpoint at 5 Years				
	Whole Cohort (N=134)		**Surgical Groups (N=96)**	
	Odds Ratio (95% CI)	**P value**	**Odds Ratio (95% CI)**	**P value**
Final Model				
Duration of diabetes < 8 years (yes vs no)	3.95 (1.46, 10.69)	0.007	4.31(1.45, 12.77)	0.008
Gastric bypass (vs sleeve)	----		1.23 (0.47, 3.20)	0.670
Gastric bypass (vs medical)	4.34 (1.11, 16.90)	0.034	----	
Sleeve (vs medical)	3.52 (0.88, 14.08)	0.075	----	
Non-significant baseline variables considered for selection				
Female (vs Males)	0.97 (0.37, 2.57)	0.953	0.76 (0.27, 2.14)	0.602
Insulin Use (yes vs no)	0.55 (0.21, 1.43)	0.547	0.71 (0.25, 1.97)	0.507
Baseline BMI (per unit BMI)	1.00 (0.88, 1.14)	0.964	0.99 (0.86, 1.14)	0.902
Baseline HbA1c (per unit HbA1c,%)	0.87 (0.64, 1.20)	0.412	0.92 (0.66, 1.27)	0.601
Baseline C-peptide (per ng/ml)	1.16 (0.86, 1.57)	0.322	1.07 (0.79, 1.45)	0.661
Baseline C-reactive protein (log), (per log increase)	1.03 (0.64, 1.62)	0.904	1.04 (0.64, 1.69)	0.875
Age (per year)	0.97 (0.92, 1.03)	0.33	0.96 (0.91, 1.02)	0.176
Baseline LDL (per mg/dl)	0.99 (0.98, 1.01)	0.46	0.99 (0.98, 1.01)	0.492
Baseline HDL (per mg/dl)	0.99 (0.96, 1.04)	0.86	0.98 (0.95, 1.03)	0.519
Baseline FPG (per mg/dl)	1.00 (0.99, 1.01)	0.62	1.00 (0.99, 1.01)	0.572
Baseline SBP (per mmHg)	1.00 (0.98, 1.03)	0.99	0.99 (0.97, 1.02)	0.846
Baseline DBP (per mmHg)	1.00 (0.96, 1.04)	0.94	0.98 (0.94, 1.03)	0.488

Methodology

A stepwise logistic model determined factors associated with the primary endpoint, defined as achieving an HbA1c ≤6% at 5 years. The odds ratios above provide the odds of having met the primary endpoint for each variable compared to its reference (in parentheses), assuming other parameters in the model are held constant. Continuous variables report a relative increase/decrease in the odds ratio *per unit* change, assuming other parameters in the model are held constant. An alpha of 0.05 was used as the entry and stay criteria into the logistic model. All variables shown above were considered for inclusion. Estimates for non-significant variables are reported after being added one at a time, but were not kept in the final model.

Table S3. Additional Secondary Endpoints

				P Value		
	Medical Therapy (N=38)	**Gastric Bypass (N=49)**	**Sleeve Gastrectomy (N=47)**	**Gastric Bypass vs. Medical Therapy**	**Sleeve Gastrectomy vs. Medical Therapy**	**Gastric Bypass vs. Sleeve Gastrectomy**
HbA1c at 5 years, median [Interquartile range]	8.0 (7.2 to 9.2)	6.9 (6.0 to 8.2)	7.1 (6.1 to 8.0)			
HbA1c ≤ 7% without diabetes medications	0	18 (36.7)	11 (23.4)	<0.001	0.001[ex]	0.15
Relapse of glycemic control†	4/5 (80.0)	8/20 (40.0)	8/16 (50.0)	0.16[ex]	0.34[ex]	0.55
Experienced weight regain >5% after 1 year	0/4 (0.0)	0/8 (0.0)	0/8 (0.0)	NA	NA	NA
Relapse of diabetes †	NA	10/20 (50.0)	5/12 (41.7)	NA	NA	0.65
Percent change in HbA1c from baseline						
> 20% Reduction HbA1c	5 (13.2)	28 (57.1)	23 (48.9)	<0.001	0.001	0.820
No Change	27 (71.1)	19 (38.8)	22 (46.8)			
> 20% Increase HbA1c	6 (15.8)	2 (4.1)	2 (4.3)			
% Change from baseline in body weight	-5.0 ± 9.9	-21.8 ± 8.3	-18.5 ± 6.6	<0.001	<0.001	0.12
Waist – cm						
Baseline	113.7 ± 9.40	116.5 ± 9.25	113.5 ± 10.35			
5 years	111.6 ± 13.09	99.4 ± 9.23	99.3 ± 9.43			
% Change	-1.3 ± 10.17	-14.7 ± 6.60	-12.2 ± 7.96	<0.001	<0.001	0.122
Within-group p-value	0.491	<0.001	<0.001			
Waist:Hip Ratio						
Baseline	1.0 ± 0.09	1.0 ± 0.07	1.0 ± 0.08			
5 years	1.0 ± 0.08	0.9 ± 0.07	0.9 ± 0.07			
%Change from baseline	2.1 ± 4.81	-2.4 ± 4.98	-2.0 ± 8.48	<0.001	0.019	0.769
Within-group p-value	0.022	0.003	0.174			
	Achieved primary endpoint	Did not achieve primary endpoint	P-value			
% Weight loss	-23.6%	-13.8%	<0.001			

† Relapse of glycemic control was defined as having met the primary end point for glycated hemoglobin of 6% or less at 1 year but not at 5 years
Relapse of diabetes was defined as having met the primary end point for glycated hemoglobin of 6% or less with the use of no antidiabetic medications at 1 year but not at 5 years. None of the patients in the medical-therapy group had a complete remission of diabetes, and thus, these patients were not evaluated for relapse.
[Ex] Indicates an exact test was performed.

An analysis of variance with treatment group as the only factor was used to analyze the change from baseline to 5 years for the following secondary endpoints: Percentage change in body weight, waist circumference and the waist:hip ratio. Unadjusted pairwise p-values are provided as well as p-values testing the hypothesis that the difference from baseline within each treatment group is different from zero. Due to inflation of type I error related to multiple testing, p-values should be interpreted with caution.

Table S4. Medication Use at Baseline and at 5 Years*

	Baseline			5 Years		
	Medical Therapy (n=38)	Gastric Bypass (n=49)	Sleeve Gastrectomy (n=47)	Medical Therapy (n=38)	Gastric Bypass (n=49)	Sleeve Gastrectomy (n=47)
Diabetes medications – no. (%)						
Biguanides	35 (92.1)	42 (85.7)	40 (85.1)	32 (84.2)	21 (42.9) †	28 (59.6) ¶
TZD's	16 (42.1)	25 (51.0)	16 (34.0)	4 (10.5)	0 (0.0) ¶	1 (2.1)
Incretin mimetics	17 (44.7)	20 (40.8)	20 (42.6)	12 (31.6)	8 (16.3)	13 (27.7)
Secretagogues	16 (42.1)	17 (34.7)	17 (36.2)	9 (23.7)	12 (24.5)	11 (23.4)
Insulin	20 (52.6)	23 (46.9)	21 (44.7)	15 (39.5)	6 (12.2) †	5 (10.6) †
Injectables‡	25 (65.8)	30 (61.2)	29 (61.7)	20 (52.6)	9 (18.4) †	10 (21.3) †
Average number of diabetes medications, mean±s.d	2.76 ± 1.12	2.61 ± 1.11	2.45 ± 1.21	2.05 ± 1.04	1.10 ± 1.19 †	1.36 ± 1.05 ¶
Number of diabetes medications – no. (%)						
0	1 (2.6)	0 (0.0)	1 (2.1)	1 (2.5)	22 (44.9) €	12 (25.5) €¥
1	4 (10.5)	10 (20.4)	11 (23.4)	12 (31.6)	9 (18.4)	14 (29.8)
2	10 (26.3)	13 (26.5)	13 (27.7)	13 (34.2)	10 (20.4)	13 (27.7)
≥3	23 (60.5)	26 (53.1)	22 (46.8)	12 (31.6)	8 (16.3)	8 (17.0)
Cardiovascular medications – no. (%)						
Lipid lowering agents	32 (84.2)	42 (85.7)	37 (78.7)	27 (71.1)	19 (38.8) †	26 (55.3)
Beta-blocker	7 (18.4)	9 (18.4)	6 (12.8)	10 (26.3)	11 (22.4)	3 (6.4) ¶δ
Calcium channel blocker	4 (10.5)	4 (8.2)	2 (4.3)	1 (2.6)	3 (6.4)	3 (6.4)
ACE-inhibitor or ARB	24 (63.2)	37 (75.5)	30 (63.8)	14 (36.8)	19 (38.3)	12 (25.5)
Diuretics	12 (31.6)	18 (36.7)	14 (29.8)	13 (34.2)	9 (18.4)	19 (40.4) δ
Anticoagulant	21 (55.3)	21 (42.9)	15 (31.9)	12 (31.6)	2 (4.1) †	10 (21.3) δ
Average number of cardiovascular medications, mean±s.d.	2.74 ± 1.22	2.78 ± 1.28	2.21 ± 1.06	2.16 ± 1.33	1.37 ± 1.35¶	1.57 ± 1.32
Number of CV medications – no. (%)						
0	0 (0.0)	3 (6.1)	1 (2.1)	4 (10.5)	17 (34.7) €	13 (27.7)
1	5 (13.2)	4 (8.2)	12 (25.5)	9 (23.7)	12 (24.5)	11 (23.4)
2	13 (34.2)	12 (24.5)	16 (34.0)	10 (26.3)	10 (20.4)	10 (21.3)
≥3	20 (52.6)	30 (61.2)	18 (38.3)	15 (39.5)	10 (20.4)	13 (27.7)

*All p-values were calculated on the basis of the 60-month data with the medical-therapy group as the comparator.
ACE, angiotensin converting enzyme; ARB, angiotensin receptor blocker. ‡ Injectables includes insulin
† P-value <0.001
¶ P-value <0.01
€ P-value <0.05 for categorical comparison to intensive medical therapy
δ P-value <0.05 for comparison between gastric bypass and sleeve gastrectomy
¥ P-value <0.05 for comparison of "No Diabetes Medications" between gastric bypass and sleeve gastrectomy

Medication data at baseline and 5 years is provided above. Categorical variables are summarized with the use of frequencies. Pearson's chi-square or a Fisher's Exact test was performed to evaluate differences in medication usage at 5 years among the 3 treatment groups. Unadjusted pairwise p-values were calculated and should be interpreted with caution due to type I error inflation associated with multiple testing.

Table S5. Renal Endpoints through 5 Years*

				P Value		
	Medical Therapy (n=37)	Gastric Bypass (n=47)	Sleeve Gastrectomy (n=45)	Gastric Bypass vs. Medical Therapy	Sleeve Gastrectomy vs. Medical Therapy	Gastric Bypass vs. Sleeve Gastrectomy
Albumin/creatinine ratio						
Baseline	6 (5.0, 12.0)	9 (5.0, 49.0)	12 (7.0, 22.0)			
Change from baseline	0.5 (-3.0, 16.0)	-1.0 (-20.1, 6.0)	-5.0 (-15.0, -1.0)			
% change from baseline	7.1 (-42.9, 105.9)	-16.7 (-79.4, 88.9)	-59.5 (-70.4, -14.3)	0.203	<0.001	0.098
Within-group p-value	0.124	0.773	<0.001			
Serum creatinine (mg/dL)						
Baseline	0.7 (0.60, 0.80)	0.7 (0.57, 0.82)	0.7 (0.58, 0.80)			
Change from baseline	0.02 (-0.06, 0.13)	0.11 (0.00, 0.22)	0.08 (0.01, 0.18)	0.127	0.077	0.959
Within-group p-value	0.111	<0.001	<0.001			
Glomerular filtration rate†						
Baseline	106 (97.0, 111.7)	110 (97.7, 119.2)	109 (96.5, 114.2)			
% change from baseline	-1.1 (-11.0, 3.3)	-7.7 (-16.2, 0.0)	-6.2 (-15.5, -0.6)	0.184	0.051	0.742
Within-group p-value	0.204	0.001	<0.001			
Albuminuria‡						
Baseline	3 (8.8)	13 (27.7)	8 (17.8)	0.036	0.335	0.259
Albuminuria status at 60 Months						
No Albuminuria	24 (70.6)	30 (63.8)	35 (77.8)	0.135	0.146	0.554
Developed Albuminuria	7 (20.6)	4 (8.5)	2 (4.4)			
Resolved Albuminuria	2 (5.9)	8 (17.0)	5 (11.1)			
Sustained Albuminuria	1 (2.9)	5 (10.6)	3 (6.7)			

*Values are median (IQR)
For skewed data, such as the albumin/creatinine ratio, the median of the percentage change is not the numerical difference between the group-level medians at baseline and at 3 years.
† calculated using the CKD-EPI formula. GFR = 141 X min (Scr/κ, 1) α X max (Scr/κ, 1)-1.209 X 0.993Age X 1.018 [if female] X 1.159 [if black]; where Scr is serum creatinine (mg/dL), κ is 0.7 for females and 0.9 for males, α is –0.329 for females and –0.411 for males, min indicates the minimum of Scr/κ or 1, and max indicates the maximum of Scr/κ or 1
‡ Albuminuria defined as urine albumin/creatinine ratio >30 mg/g

Descriptive data on the renal and microvascular endpoints at 5 years are provided above. Non-parametric methods using pairwise 2-sample Wilcoxon test statistics were used to generate the unadjusted p-values. The Wilcoxon signed-rank test was used to test for differences from baseline. Due to inflation of type I error related to multiple testing, p-values should be interpreted with caution.

Table S6a. Retinopathy at Baseline and at 5 Years

	Medical Therapy (n=25)	Gastric Bypass (n=42)	Sleeve Gastrectomy (n=36)	P Value
Retinopathy at Baseline				
None	22 (88.0)	33 (78.6)	30 (83.3)	0.65
NPDR - Mild	2 (8.0)	8 (19.0)	3 (8.3)	
NPDR - Moderate	1 (4.0)	1 (2.4)	1 (2.8)	
NPDR - Severe	0 (0.0)	0 (0.0)	1 (2.8)	
PDR - Non High Risk	0 (0.0)	0 (0.0)	1 (2.8)	
Any Retinopathy at Baseline	3 (12.0)	9 (21.4)	6 (16.7)	0.61
Retinopathy at 5 years – Number with available data				
None	23 (92.0)	35 (83.3)	31 (86.1)	0.31
NPDR - Mild	1 (4.0)	5 (11.9)	3 (8.3)	
NPDR - Moderate	0 (0.0)	2 (4.8)	0 (0.0)	
NPDR - Severe	1 (4.0)	0 (0.0)	0 (0.0)	
PDR - Non High Risk	0 (0.0)	0 (0.0)	2 (5.6)	
Any Retinopathy at 5 Years	2 (8.0)	6 (16.7)	4 (13.9)	0.67
p-value for baseline vs 5 years	*0.33*	*0.67*	*0.20*	

NPDR=Non-proliferative diabetic retinopathy. PDR=Proliferative diabetic retinopathy. The Chi-square statistic was used to compare the categories of retinopathy at baseline and 5years across the 3 treatment strategies. The Bowker's test of symmetry was used test that proportions were symmetrical between baseline and 5 years across the 3 groups.

Table S6b. Visual Acuity at Baseline and at 5 Years

	Baseline		**5 Years**	
	LogMar VA OD (Approximate Snellen)	**LogMar VA OS (Approximate Snellen)**	**LogMar VA OD (Approximate Snellen)**	**LogMar VA OS (Approximate Snellen)**
Non-Surgery (n=25)	0.037 ± 0.0813 (20/21.8)	0.018 ± 0.0515 (20/20.8)	-.000 ± 0.0979 (20/20)	0.013 ± 0.1226 (20/20.6)
Gastric Bypass (n=42)	0.021 ± 0.0587 (20/21)	0.020 ± 0.0507 (20/20.9)	0.045 ± 0.1457 (20/22.2)	0.030 ± 0.0876 (20/21.4)
Sleeve Gastrectomy (n=36)	0.021 ± 0.0452 (20/21)	0.027 ± 0.0639 (20/21.3)	0.048 ± 0.1603 (20/22.3)	0.049 ± 0.1288 (20/22.4)
Total (n=103)	0.025 ± 0.0608 (20/21.2)	0.022 ± 0.0554 (20/21)	0.035 ± 0.1415 (20/21.7)	0.032 ± 0.1118 (20/21.5)
p-value	0.397	0.843	0.144	0.254

The LogMar values across treatment groups were tested using the Kruskal-Wallis test at baseline and at 5 years.

Table S6c. Macular Edema at Baseline and 5 Years

	Medical Therapy (n=25)	Gastric Bypass (n=42)	Sleeve Gastrectomy (n=35)	P Value
Any macular edema at baseline	1 (4.0)	3 (7.1)	1 (2.9)	0.67
Any macular edema at 5 years	1 (4.0)	0	0	0.21

Any macular edema includes non-clinically significant and clinically significant macular edema. The chi-square statistic was used to assess macular edema at baseline and 5 years across the 3 treatment strategies.

Table S7. RAND-36 Questionnaire Responses at Baseline and at 5 Years

	Intensive Medical Therapy (n=24)	Gastric Bypass (n=39)	Sleeve Gastrectomy (n=37)	Gastric Bypass vs. Medical Therapy	Sleeve Gastrectomy vs. Medical Therapy	Gastric Bypass vs. Sleeve Gastrectomy
Physical Health Components						
Physical Functioning						
Baseline	75.1 ± 19.66	78.5 ± 16.35	79.5 ± 15.79			
Change from baseline	1.6 ± 17.65	8.1 ± 18.28	7.2 ± 29.19	0.172	0.402	0.876
p-value for change from baseline	*0.39*	*0.002*	*0.01*			
Role Limitations due to Physical Health						
Baseline	76.0 ± 39.34	84.8 ± 28.15	79.7 ± 32.19			
Change from baseline	-10.4 ± 55.62	4.1 ± 36.87	0.7 ± 44.59	0.218	0.393	0.719
p-value for change from baseline	0.34	0.51	0.99			
Bodily Pain						
Baseline	82.7 ± 14.76	80.5 ± 19.14	79.5 ± 17.82			
Change from baseline	-16.6 ± 25.35	-2.4 ± 25.20	0.5 ± 20.55	0.037	0.005	0.580
p-value for change from baseline	*0.01*	*0.77*	*0.87*			
General Health						
Baseline	54.7 ± 19.80	55.1 ± 15.83	52.3 ± 20.33			
Change from baseline	0.3 ± 16.04	17.4 ± 20.02	16.0 ± 22.20	0.001	0.004	0.782
p-value for change from baseline	*0.92*	*<0.001*	*<0.001*			
Mental Health Components						
Role Limitations due to Emotional Problems						
Baseline	88.9 ± 27.22	94.0 ± 20.05	89.2 ± 26.12			
Change from baseline	-14.5 ± 38.70	-3.4 ± 23.93	-12.0 ± 40.75	0.168	0.819	0.263
p-value for change from baseline	*0.10*	*0.47*	*0.82*			
Energy/Fatigue						
Baseline	49.6 ± 20.00	54.9 ± 18.90	46.1 ± 24.04			
Change from baseline	5.3 ± 25.63	11.4 ± 19.40	17.6 ± 27.15	0.299	0.083	0.261
p-value for change from baseline	*0.32*	*0.001*	*0.001*			
Emotional Well Being						
Baseline	83.6 ± 10.48	85.4 ± 8.71	79.1 ± 14.82			
Change from baseline	-8.8 ± 18.13	-5.9 ± 15.43	0.3 ± 21.40	0.502	0.090	0.158
p-value for change from baseline	*0.04*	*0.03*	*0.62*			
Social Functioning						
Baseline	88.5 ± 16.86	88.5 ± 15.81	82.1 ± 21.15			
Change from baseline	-9.4 ± 28.13	-0.3 ± 20.30	3.7 ± 28.70	0.150	0.085	0.485
p-value for change from baseline	*0.13*	*0.99*	*0.44*			

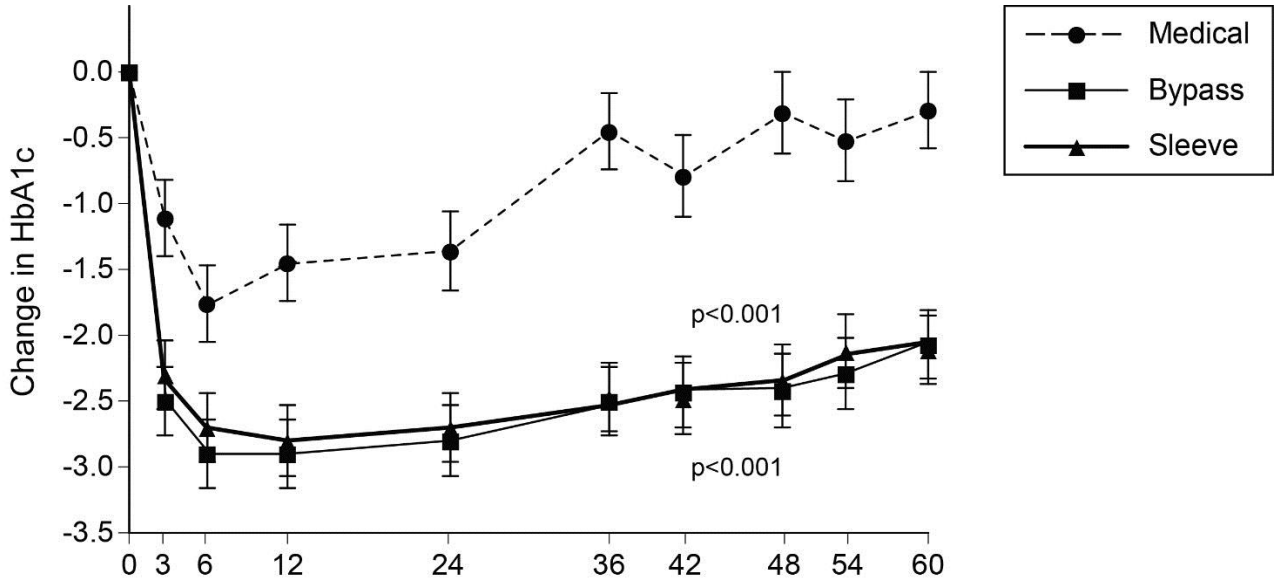
Each item of the RAND-36 data are summarized using the patient's original response. There was no imputation for non-response. P-values for the pairwise comparisons were generated using an analysis of variance with treatment group as a factor. Changes from baseline within each treatment group were evaluated with a paired t-test. Given the multiple testing performed, p-values should not be considered conclusive and should be interpreted with caution.

Table S8. Additional Imputation Results

				P Value		
	Medical Therapy (N=50)	**Gastric Bypass (N=50)**	**Sleeve Gastrectomy (N=50)**	**Gastric Bypass vs. Medical Therapy**	**Sleeve Gastrectomy vs. Medical Therapy**	**Gastric Bypass vs. Sleeve Gastrectomy**
Minimum value imputation	3 (6.0)	15 (30.0)	13 (26.0)	0.002	0.006	0.656
Mean value imputation	3 (6.0)	14 (28.0)	11 (22.0)	0.003	0.02	0.488

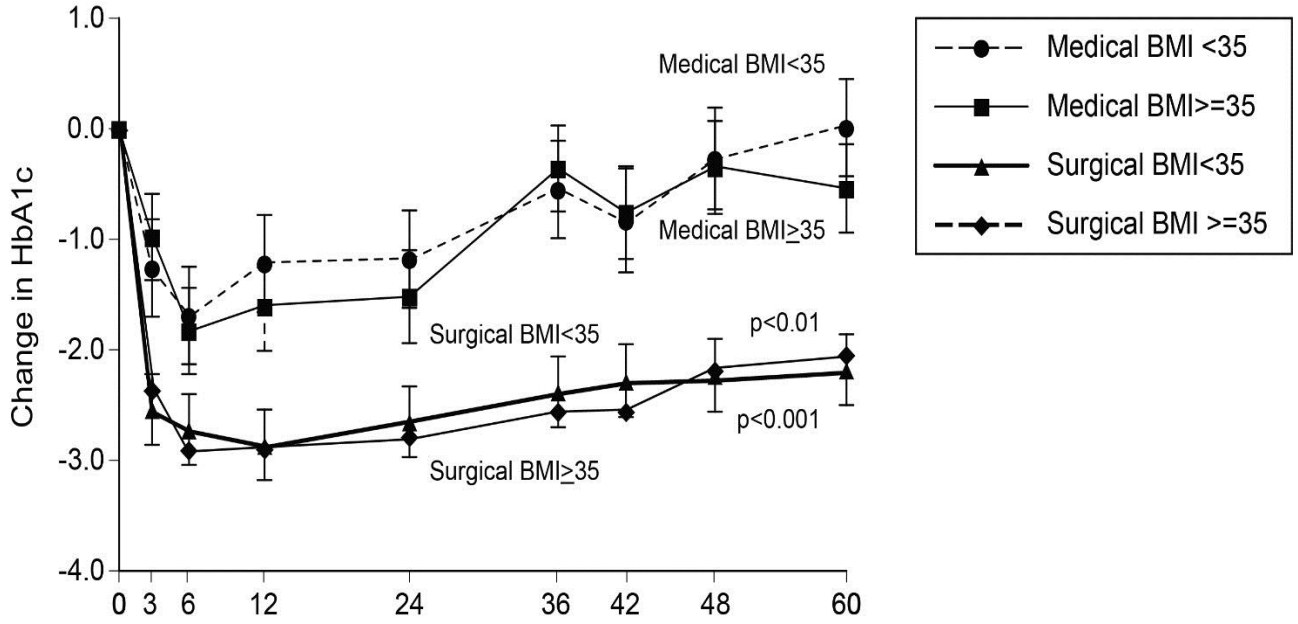
Two additional simple substitution methods were used for imputing the primary endpoint in patients where data was unavailable: 1) the minimum value across all study visits (including baseline) was used to impute the final value at the end of study. The dichotomous primary endpoint (HbA1c ≤ 6%) was then computed using this imputed value, and 2) the mean 5 year HbA1c value for each treatment group was used to impute the final value at the end of the study. Again, the dichotomous primary endpoint (HbA1c ≤ 6%) was computed using this imputed value. Unadjusted p-values for each pairwise comparison were generated using the chi-square statistic and should be interpreted with caution due to inflation of type I error related to multiple testing.

Figure S1. Change in Glycated Hemoglobin



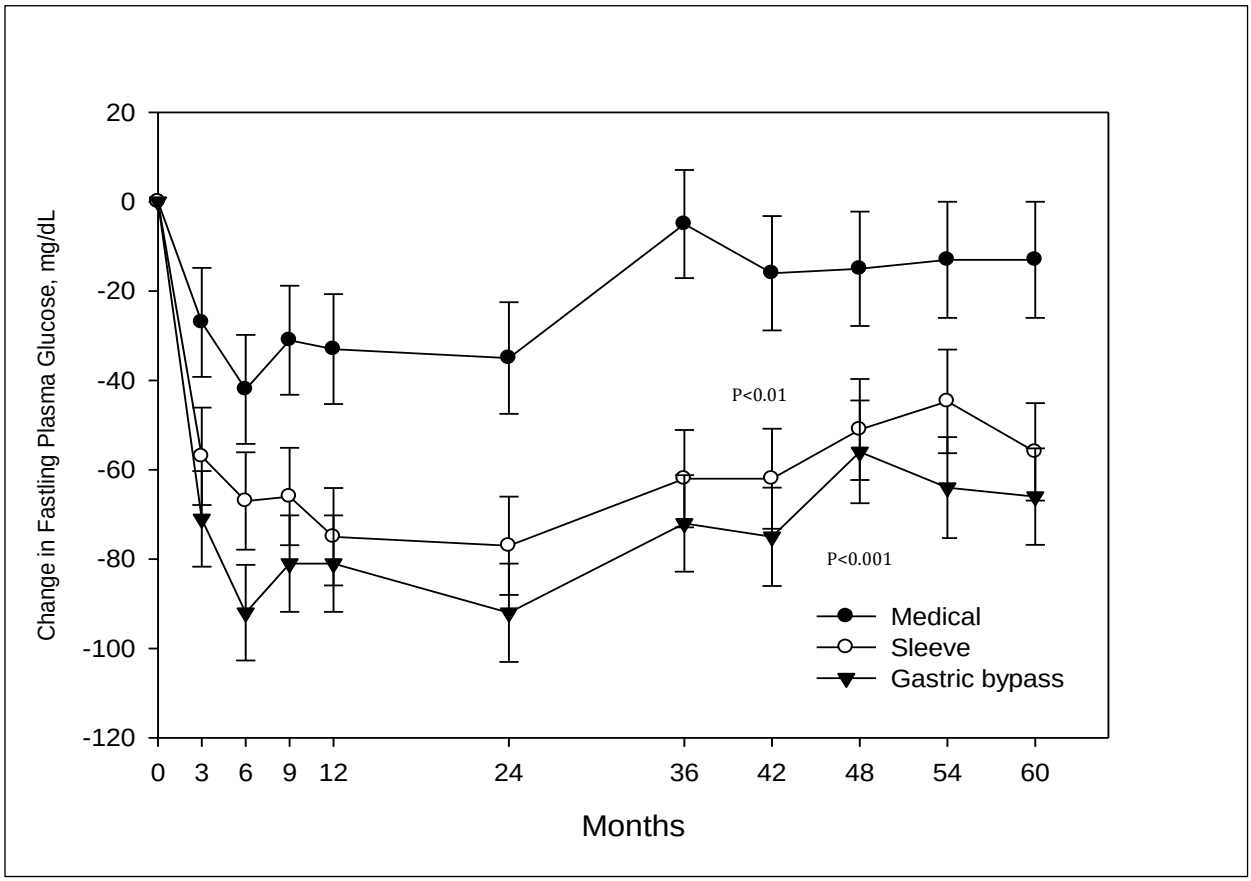
Value at Visit	Baseline	Month 12	Month 24	Month 36	Month 48	Month 60
Medical	8.8 (8.6)	7.3 (6.8)	7.5 (7.2)	8.4 (7.7)	8.6 (8.2)	8.5 (8.0)
Gastric Bypass	9.3 (9.4)	6.4 (6.2)	6.5 (6.4)	6.8 (6.6)	6.8 (6.8)	7.3 (6.9)
Sleeve	9.5 (8.9)	6.7 (6.4)	6.8 (6.8)	7.0 (6.7)	7.1 (6.6)	7.4 (7.2)

Figure S2. Change in Glycated Hemoglobin According to Body Mass Index



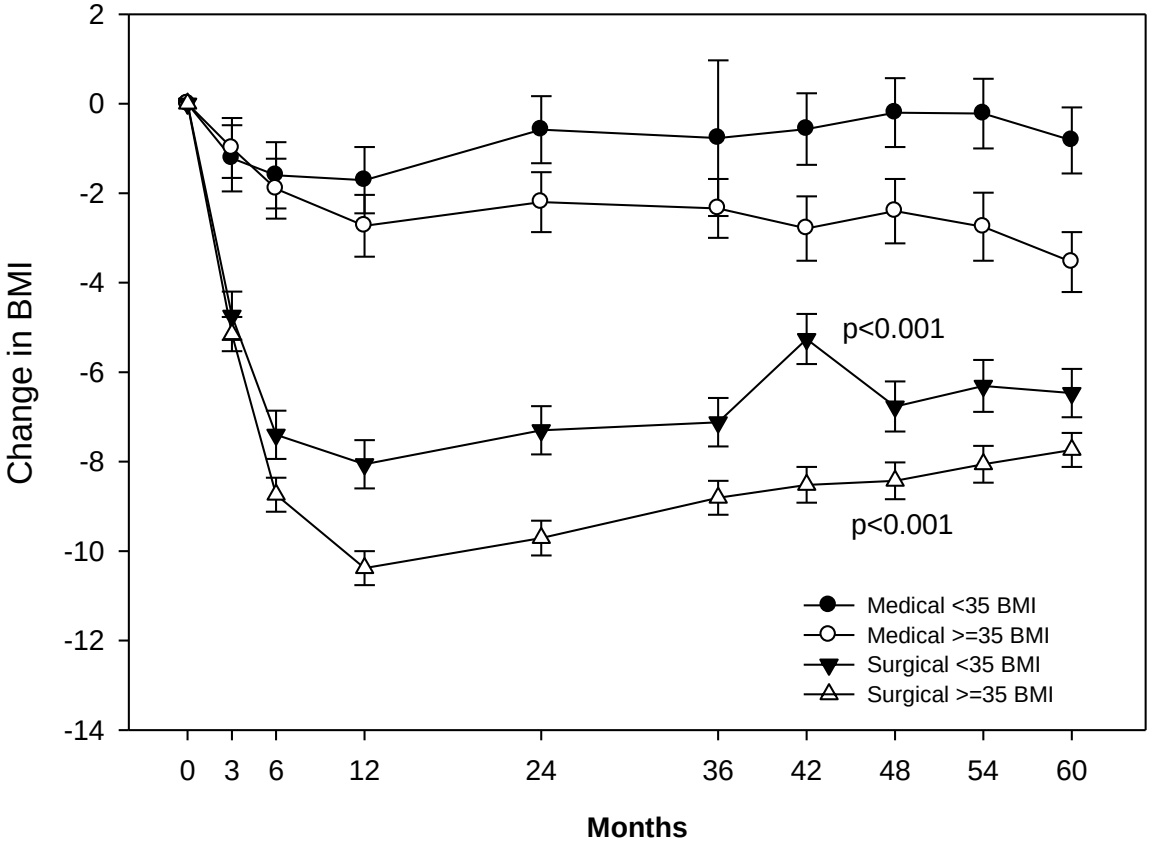
Value at Visit	Baseline	Month 12	Month 24	Month 36	Month 48	Month 60
Medical <35 (N=17)	8.8 (8.9)	7.5 (6.9)	7.7 (7.4)	8.2 (7.9)	8.8 (8.6)	8.8 (8.0)
Medical ≥35 (N=21)	8.9 (8.5)	7.2 (6.5)	7.3 (6.8)	8.5 (7.1)	8.5 (8.2)	8.3 (8.0)
Surgical <35 (N=32)	9.5 (9.1)	6.6 (6.7)	6.8 (6.8)	7.1 (6.7)	7.2 (6.8)	7.3 (7.1)
Surgical ≥35 (N=64)	9.4 (9.2)	6.5 (6.2)	6.6 (6.4)	6.8 (6.6)	6.8 (6.5)	7.3 (7.1)

Figure S3. Change in Fasting Plasma Glucose



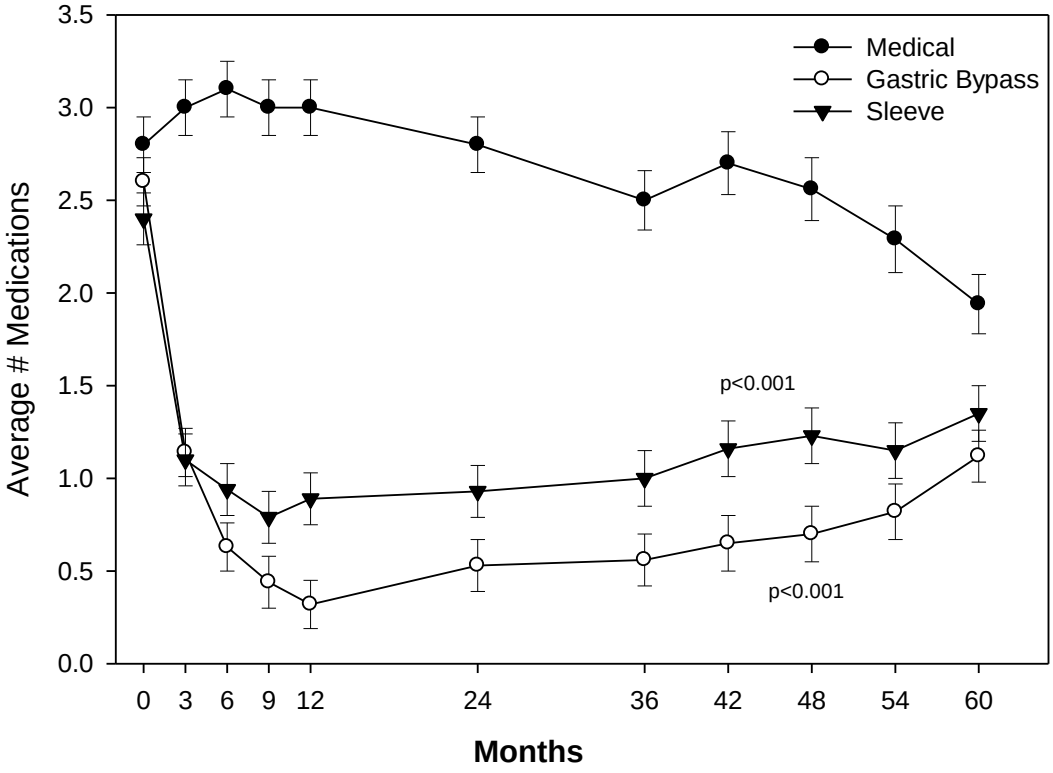
P-value is for each surgical group compared to medical therapy

Figure S4. Change in Body Mass Index: Medical vs Surgical by BMI Subgroup



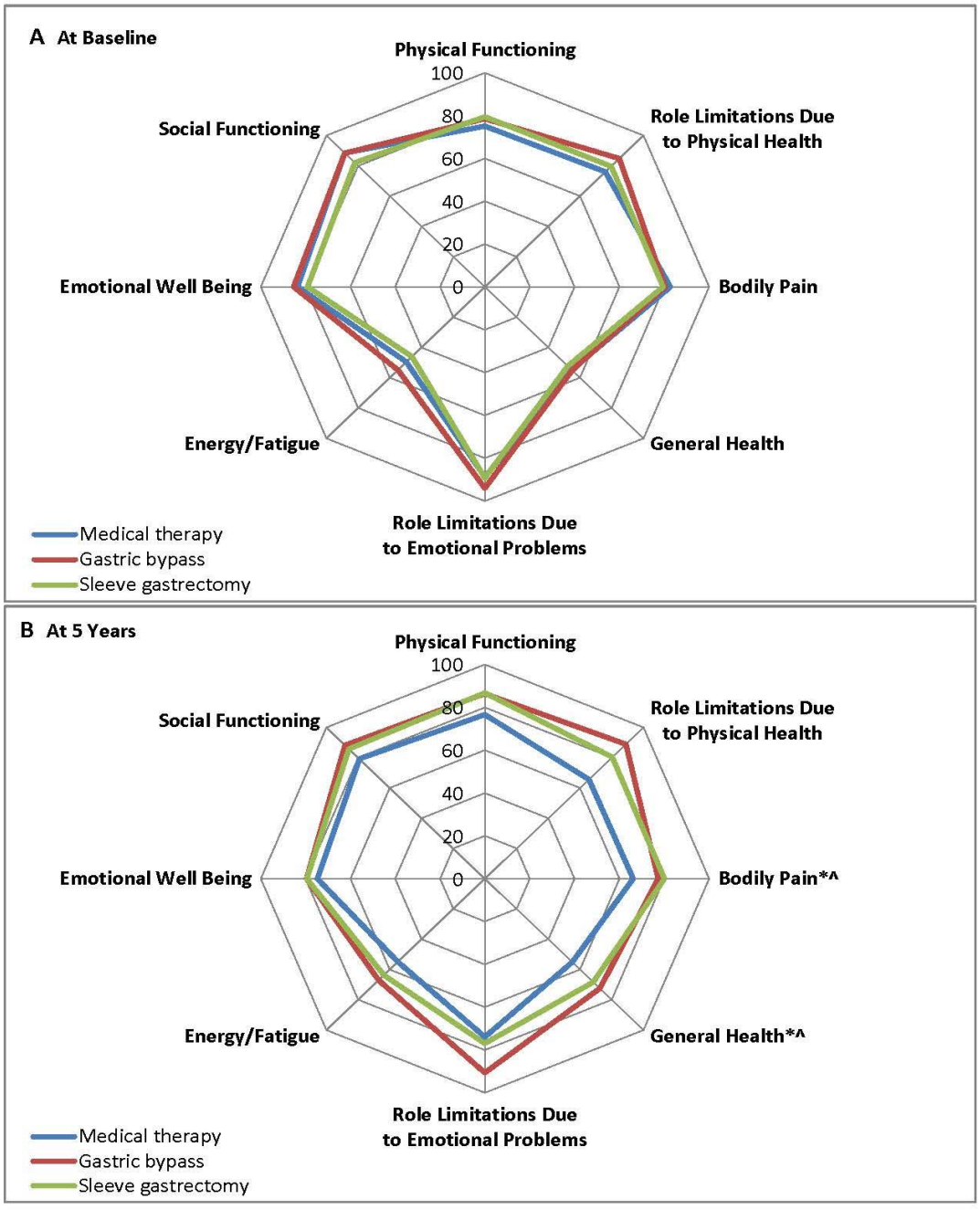
P-value is for surgical vs intensive medical therapy for each BMI subgroup

Figure S5. Average Number of Diabetes Medications



P-values are for surgical groups compared to intensive medical treatment

Figure S6. Polar Chart Scores for Quality of Life at Baseline and 5 Years after Randomization



*p<0.05 for change from baseline between gastric-bypass group and medical therapy group
^p<0.05 for change from baseline between sleeve-gastrectomy and medical therapy group