PCORnet Bariatric Study Analytic Plan

Prepared by: David Arterburn, Jason Block, Lingling Li, Darren Toh, and Rob Wellman on behalf of the PCORnet Bariatric Study

Aim 4-6 additions prepared by: David Arterburn, Yates Coley, and Rob Wellman on behalf of the PCORnet Bariatric Study

Submitted to PCORI	*November 25, 2019*

PCORnet Bariatric Study Analytic Plan

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The PCORnet Bariatric Study ("the study") team has prepared a detailed analysis plan for accomplishing the main study aims. This plan was developed by the Methods Core with feedback from the Scientific Core, Executive Bariatric Stakeholder Advisory Group, and the Clinical Data Research Network (CDRN) Bariatric Principal Investigators.

The study seeks to answer three main scientific questions:

Aim 1: To what extent does weight loss and weight regain differ across the three most common bariatric surgical procedures in the United States – Roux-en-Y Gastric Bypass (RYGB), Adjustable Gastric Banding (AGB), and Sleeve Gastrectomy (SG) – at 1, 3, and 5 years after surgery?

Aim 2: To what extent do the three most common bariatric procedures in the United States differ with respect to diabetes status at 1, 3, and 5 years after surgery?

Aim 3: What is the frequency of major adverse events for the three most common bariatric procedures in the United States at 1, 3, and 5 years?

Aim 4: To what extent do bariatric surgery outcomes explored in Aims 1-3 differ with respect to baseline **depression diagnosis** at 1, 3, and 5 years after surgery?

Aim 5: To what extent do bariatric surgery outcomes explored in Aims 1-3 differ with respect to **race and ethnicity** at 1, 3, and 5 years after surgery?

Aim 6: To what extent do bariatric surgery outcomes explored in Aims 1-3 differ with respect to **baseline age 65+ versus <65** at 1, 3, and 5 years after surgery? **(Aim 6 will only be conducted if time and resources allow this work)**

Population: The study will include adults, children, and adolescents less than 80 years old at time of surgery who had one of the three most common procedures in the United States (RYGB, AGB, or SG) during the years 2005 to 2015. To be eligible for the study all patients will also need to have a Body Mass Index (BMI) measurement in the year prior to surgery that is at least 35 kg/m^2.

Data: All data necessary for accomplishing the main scientific aims of this study will be derived from the PCORnet Common Data Model (CDM). In the plan below we have specified each CDM table that will be accessed as well as the key variables that will be examined. Although all of the tables described are necessary to complete the study, of greatest interest are the procedures, diagnosis, vitals, prescribing, dispensing, and death tables. All data for this study will be abstracted from the CDM tables at each participating health care site and then sent using secured file transfer methods to the data coordinating center at Harvard Pilgrim and to Kaiser Permanente Washington Health Research Institute for analysis.

Analyses: The "primary analyses" will address the main study questions outlined above using individual-level patient data. We will conduct three pair-wise comparisons for each study aim – comparing AGB versus RYGB, SG versus RYGB, and AGB vs. SG. To address potential confounding bias in each comparison, we will first use a logistic regression model to estimate the propensity score (PS), which is

defined as the probability of receiving a treatment of interest (e.g., RYGB) in each pairwise comparison given the potential confounders variables plus calendar year. We will conduct sensitivity analyses to assess for changes in outcomes by calendar year, and we will use a multiple imputation approach to address missing outcome information. Finally, for each aim, we will seek to identify heterogeneity of treatment effects, with are any differences in the effects of the bariatric procedures across key subgroups (e.g., race/ethnicity).

Methodology Standards: Throughout the document, references are made to PCORI's Methodology Standards (e.g., [RQ-1]). A description of these standards can be found here on the PCORI website

AIMS AND ANALYTIC PLAN SYNOPSIS

This section provides an overview of the primary and secondary analyses for each aim, as well as their primary and secondary outcomes. An overview of sensitivity analyses and analyses of heterogeneity of treatment effects is also provided for each aim.

A. AIM 1 (WEIGHT LOSS AND WEIGHT REGAIN)

To what extent does weight loss and weight regain differ across the three bariatric surgical (RYGB, AGB, and SG) at 1, 3, and 5 years?

Primary Analysis: Complete three pairwise comparisons (RYGB vs. AGB; RYGB vs. SG; AGB vs. SG) for our primary and secondary outcomes and all sensitivity/heterogeneity analyses on adults ≥20 years of age using *individual-level data with multiple imputation for missing weight data after surgery*

Secondary Analysis: Complete three pairwise comparisons for our primary and secondary outcomes and all sensitivity/heterogeneity analyses using *distributed data analysis approach*

Primary Outcome: percentage change in weight (in kg) at 1, 3, and 5 years compared to baseline

Secondary Outcomes: 1) weight regain at 3 and 5 years; estimated as percent regain from the maximum weight (in kg) loss in the first 2 years; 2) a post-operative body weight that is <5% lower than the pre-surgical weight at 1, 3, or 5 years after bariatric surgery; and 3) proportion achieving >5%, >10%, >20%, and >30% weight loss at 1, 3, and 5 years

Sensitivity Analyses (for primary outcome): 1) Calendar Year analysis: look at the effects of these procedures on weight changes by calendar year

Heterogeneity Analyses (for primary outcome): 1) Complete all analyses on adolescents <20 years of age; 2) Compare comorbidity groups (e.g., baseline diabetes) or comorbidity index score quartiles 3) Compare baseline BMI (≥50 vs <50 kg/m^2), 4) Compare adults <65 years and 65+ years; 5) Compare race/ethnicity (e.g., non-Hispanic black, Hispanic, non-Hispanic white), 6) Gender (male vs. female)

B. AIM 2 (IMPROVEMENTS IN DIABETES RISK)

To what extent do these bariatric procedures differ on improvements in diabetes risk at 1, 3, and 5 years?

Primary Analysis: Complete three pairwise comparisons (RYGB vs. AGB; RYGB vs. SG; AGB vs. SG) for our primary and secondary outcomes and all sensitivity/heterogeneity analyses on adults ≥20 years of age using *individual-level data with multiple imputation for missing weight data after surgery*

Secondary Analysis: Complete three pairwise comparisons for our primary and secondary outcomes and all sensitivity/heterogeneity analyses using *distributed data analysis approach*

Primary Outcomes: 1) Rate of diabetes remission (HbA1c <6.5% off diabetes medications ≥90 days);

Secondary Outcomes: 1) Rate of relapse of diabetes after initial remission (restart of medication or HbA1c ≥6.5%), and 2) Change in HbA1c at 1, 3, and 5 years;

Sensitivity Analyses (for primary outcome): 1) Calendar Year analysis: look at the effects of these procedures on diabetes remission by calendar year

Heterogeneity Analyses (for primary outcome): 1) Complete all analyses on adolescents ≤20 years of age; 2) Compare individuals with poorly controlled diabetes (HbA1c ≥7%) vs. controlled diabetes at baseline, 3) Compare insulin use vs. no insulin at baseline;) 4) Compare comorbidity groups or comorbidity index score quartiles; 5) Compare baseline BMI (≥50 vs <50 kg/m2), 6) Compare race/ethnicity (e.g., non-Hispanic black, Hispanic, non-Hispanic white), 7) Compare adults <65 years and 65+ years; 8) Gender (male vs. female).

C. AIM 3 (MAJOR ADVERSE EVENTS)

What is the frequency of major adverse events following these three different bariatric surgical procedures at 1, 3, and 5 years?

Primary Analysis: Complete three pairwise comparisons (RYGB vs. AGB; RYGB vs. SG; AGB vs. SG) for our primary and secondary outcomes and all sensitivity/heterogeneity analyses on adults >20 years of age using *individual-level data.*

Primary Outcomes: 1) Time until reoperation or re-intervention.

Primary Comparison: Hazard ratio of reoperation/re-intervention for RYGB vs. SG, RYGB vs. AGB, and SG vs. AGB.

Sensitivity Analysis (for primary outcome): 1) include endoscopic procedures in re-intervention definition.

Heterogeneity Analyses (for primary outcome): 1) Compare adults <65 years and 65+ years; 2) Compare Elixhauser comorbidity index score groups (<0; 0; >0); 3) Compare baseline BMI (three groups: BMI ≥60, 60>BMI>=50, BMI <50 kg/m^2); 4) Compare race/ethnicity (e.g., non-Hispanic black, Hispanic, non-Hispanic white); 5) Gender (male vs. female); 6) Diabetes vs no Diabetes.

Secondary Outcomes: 1) Time until mortality (all-cause); 2)Occurrence of composite adverse event (AE) metric derived from the bariatric literature (censored at 30 days); and 3) Time until rehospitalization.

Secondary Comparisons: 1) hazard ratio (HR) for all-cause mortality; 2) odds ratio (OR) for 30-day composite AE; and, 3) HR for rehospitalization for RYGB vs. SG, RYGB vs. AGB, and SG vs. AGB.

D. AIM 4 (SURGERY OUTCOMES AND BASELINE DEPRESSION)

To what extent do bariatric surgery outcomes explored in Aims 1-3 differ across three bariatric procedures (AGB, RYGB, and SG) with respect to baseline **depression diagnosis** at 1, 3, and 5 years after surgery?

Analyses: Complete heterogeneity in treatment effect (HTE) analyses for three pairwise comparisons (RYGB vs. AGB; RYGB vs. SG; AGB vs. SG) for primary and secondary outcomes from Aims 1-3.

Outcomes (and comparisons):

(1) Percent total weight loss at 1, 3, and 5 years (difference in percent total weight loss)
(2) Proportion of patients with >5%, 10%, 20%, and 30% total weight loss (odds ratio)
(3) Rate of diabetes remission (hazards ratio)
(4) Rate of relapse after initial remission (hazards ratio)
(5) Change in HbA1c at 1, 3, and 5 years (difference in change in HbA1c)
(6) Time until reoperation or reintervention with and without endoscopy (hazards ratio)
(7) Time until all-cause mortality (hazards ratio)
(8) Time until hospitalization (hazards ratio)
(9) Proportion of patients with 30-day composite adverse events (odds ratio)

E. AIM 5 (SURGERY OUTCOMES AND RACE/ETHNICITY)

To what extent do bariatric surgery outcomes explored in Aims 1-3 differ across three bariatric procedures (AGB, RYGB, and SG) with respect to baseline **depression diagnosis** at 1, 3, and 5 years after surgery?

Analyses: Complete heterogeneity in treatment effect (HTE) analyses for three pairwise comparisons (RYGB vs. AGB; RYGB vs. SG; AGB vs. SG) for primary and secondary outcomes from Aims 1-3.

Outcomes (and comparisons):

(10)Percent total weight loss at 1, 3, and 5 years (difference in percent total weight loss)
(11)Proportion of patients with >5%, 10%, 20%, and 30% total weight loss (odds ratio)
(12)Rate of diabetes remission (hazards ratio)
(13)Rate of relapse after initial remission (hazards ratio)
(14)Change in HbA1c at 1, 3, and 5 years (difference in change in HbA1c)
(15)Time until reoperation or reintervention with and without endoscopy (hazards ratio)
(16)Time until all-cause mortality (hazards ratio)
(17)Time until hospitalization (hazards ratio)
(18)Proportion of patients with 30-day composite adverse events (odds ratio)

F. AIM 6 (SURGERY OUTCOMES AND MEDICARE AGE)

To what extent do bariatric surgery outcomes explored in Aims 1-3 differ across three bariatric procedures (AGB, RYGB, and SG) with respect to baseline **age 65+ versus <65 years** at 1, 3, and 5 years after surgery?

Analyses: Complete heterogeneity in treatment effect (HTE) analyses for three pairwise comparisons (RYGB vs. AGB; RYGB vs. SG; AGB vs. SG) for primary and secondary outcomes from Aims 1-3.

Outcomes (and comparisons):

(19)Percent total weight loss at 1, 3, and 5 years (difference in percent total weight loss)
(20)Proportion of patients with >5%, 10%, 20%, and 30% total weight loss (odds ratio)
(21)Rate of diabetes remission (hazards ratio)
(22)Rate of relapse after initial remission (hazards ratio)
(23)Change in HbA1c at 1, 3, and 5 years (difference in change in HbA1c)
(24)Time until reoperation or reintervention with and without endoscopy (hazards ratio)
(25)Time until all-cause mortality (hazards ratio)

(26)Time until hospitalization (hazards ratio)
(27)Proportion of patients with 30-day composite adverse events (odds ratio)

DATA SOURCES

All data for the main scientific aims of this study will be derived from the PCORnet Common Data Model (CDM). The specific CDM tables and variables that are important to this study have been specified in Table 1 below.

DEFINING THE STUDY SAMPLE

A. INCLUSION CRITERIA

1. Adults and children ages 12 through9 years at time of surgery
    a. Note that we will conduct separate primary analyses for adults (20-79 years) and adolescents, age 12 - 20 years; the cut point of <20 years was chosen to be consistent with the largest prior study of adolescent bariatric surgery, the NIH-funded Teen Longitudinal Assessment of Bariatric Surgery (Teen-LABS) study. [1]
2. Had a primary (not revision) bariatric procedure from years 2005-2015 of one of three types:
    a. Roux-en-y gastric bypass (RYGB)
    b. Adjustable gastric banding (AGB)
    c. Sleeve gastrectomy (SG)
        i. Codes for these procedures are provided in Table 2 below
3. Have a BMI measurement in the year prior to surgery that is ≥35 kg/m^2 for adults and adolescents.
    a. Note that we will conduct sensitivity analyses to examine our study outcomes among patients who were missing a BMI in electronic databases in the year before surgery

B. EXCLUSION CRITERIA

1. Their first bariatric procedure during the study period is a revision or an uncommon bariatric procedure (see Table 3 below). (Note that patients who have a primary (not revision) bariatric procedure but then subsequently go on to have a revision procedure days or years after their primary procedure will be included in the study.)
2. Those who have multiple bariatric procedures coded on the same day – unable to determine procedure type (see procedure assignment algorithm following Table 2 below)
3. Those whose bariatric procedure occurs the same day as an Emergency Department visit – likely the procedure is a revision and not initial
4. Those who have a non-inpatient or non-ambulatory encounter with a bariatric code.
5. Exclude those with Fundoplasty procedure in year before bariatric procedure.
6. Excluding those who have 1+ gastrointestinal cancer diagnosis codes in year before their bariatric procedure: (ICD-9 150.x, 151.x, 152.x, 157.x, 159.x, 171.5, 209.0x, 209.1x, 209.20, 209.23, 209.25, 209.26, 209.27, 209.29, 209.40-209.43, 209.5x, 209.60, 209.63, 209.65, 209.66, 209.67, 209.69, 211.x, 215.5, 238.1 and 199.x),
7. Those with no BMI measures in year before surgery

8. Baseline information on patient sex is missing or is a category other than “M” or “F” (for example, “U”).

C. ADDITIONAL AIM 2 INCLUSION CRITERIA

For these analyses, eligible patients must have uncontrolled or medication-controlled diabetes at the time of surgery.

1. Hemoglobin A1c (HbA1c) ≥6.5 % at the most recent measurement prior to surgery, *or*
2. Current prescription for diabetes medication at the time of surgery with the most recent HbA1c <6.5%
    a. An Insulin and Oral Diabetes Medications List can be found below in Table 4.

Note that Patients taking only metformin will be excluded unless they also have an ICD-9 code for Diabetes (250.x) or have HbA1c >=6.5% in the year prior to surgery. This includes patients with polycystic ovarian syndrome (PCOS) [2/26/16].

D. ADDITIONAL AIM 3 INCLUSION CRITERIA

1. For these analyses, eligible patients must be linked to relevant data sources in order to be included in the Aim 3 analysis cohort:
    - State or national death index (mortality outcomes)
    - Payer data/insurance claims (for AE outcomes)

2. Patients must be aged 20 through 79 years of age at time of surgery.
2. Exclude any patient with >=365 inpatient hospitalization days in the year prior to surgery.
3. Exclude any patient without male or female sex indicated.

E. AIMS 4-6 INCLUSION CRITERIA

For these analyses, inclusion criteria for Aims 1, 2, and 3 will be used for the outcomes from those aims. CDRNs will be restricted to those who agree to additional data analyses

F. ADDITIONAL AIM 5 INCLUSION CRITERIA

People with missing information for race and ethnicity will also be excluded.

G. AIM 4 AND 6 INCLUSION CRITERIA

No additional inclusion or exclusion criteria above and beyond our original analyses.

DEFINING THE STUDY OUTCOMES

Study outcomes are defined below in Table 5.

ANALYTIC PLANS BY AIM

Descriptive analysis: The study team will develop a distributed program to create the study cohort based on the eligibility criteria described above. The distributed program will also compare the baseline characteristics of the study cohort with patients who are otherwise eligible for the study but have missing information in key baseline covariates **[CI-1, CI-2]** (e.g., baseline BMI, and HbA1c among patients with diabetes). The comparison will allow us to assess the representativeness of the study cohort and help explore possible selection bias that can later be addressed in the analysis and interpretation of the study results. This descriptive analysis is different from the study data characterization that will be done prior.

We will characterize and compare eligible patients who undergo the three bariatric procedures in detail. Specifically, we will compare the distributions and missingness of the baseline covariates, the length of follow-up and missingness in post-surgery outcome measures (e.g., changes in BMI and HbA1c among patients with diabetes) **[IR-1]**. For Aim 5, we will examine differences between patients with and without race and ethnicity information.

A key consideration for observational comparative effectiveness studies is to include only patients who are eligible to receive the alternative treatments being compared. We will assess such "empirical equipoise" by examining the degree of overlap in the distributions of the estimated propensity scores for each of the three pairwise comparisons (described in detail below) **[CI-5]**. [2] The comparisons will help us identify patients who may have certain characteristics or contraindications that disqualify them from receiving one or more of the three studied bariatric procedures in typical clinical practice. We will restrict the study cohort to patients in the overlapping region of the propensity score distributions as appropriate.

A. AIM 1 (WEIGHT LOSS AND WEIGHT REGAIN)

1. Primary analysis

The primary analysis will include only patients who meet the eligibility criteria for the study and key baseline covariates (age, sex, and baseline BMI). Table 5 below shows important baseline covariates.

Continuous variables include age, baseline BMI, HbA1c, hospital length of stay during the year prior to surgery, and comorbidity score. Categorical variables include race/ethnicity, insulin use, oral hypoglycemic medication use, CDRN, node site, and year of surgery. We will aggregate comorbidity data into a score derived by combining conditions included in the Charlson and Elixhauser measures that has been externally validated among Medicare enrollees and is predictive of 30-day, 90-day, 180-day and 360-day mortality.[3] This comorbidity score can be efficiently calculated using code from a Mini-Sentinel tool. It includes age, sex, and 37 unique health conditions including both physical and mental health comorbidities. Smoking will be operationalized as current, never, or other.

The outcome of interest will be percentage change in body weight in kilograms (from baseline) at the end of 1 year, 3 years, and 5 years after surgery **[CI-3]**. As it is a continuous variable, we will inspect the distribution of the variable and transform it as appropriate. We will conduct three pair-wise comparisons – AGB versus RYGB, SG versus RYGB, and AGB vs. SG. For each comparison, we will first use a logistic regression model to estimate the propensity score (PS), which is defined as the probability of receiving the treatment of interest (which can be chosen arbitrarily) in each pairwise comparison given the potential confounders listed in Table 6 plus calendar year. The PS model will be fit at each site. We

will then assess covariate overlap by comparing the PS distributions by surgery group. If patients from different surgery groups differ systematically, we will consider trimming [4] to exclude patients that are very unlikely to be considered for the other surgery option and thus should not be included for comparative effectiveness assessment **[CI-5]**. We will consider using the linear mixed effects model or the generalized estimating equation model with a link to estimate the effect of surgery on percent weight changes. We will consider different PS-based analyses to adjust for confounding bias, including PS regression, matching, and inverse probability treatment weighting. For PS regression, we create a categorical variable using site-specific PS percentiles (e.g., deciles) and adjust for it in the outcome regression model. In PS matching, for each exposed subject, we use the nearest matching algorithm to find a comparable control whose PS is close to the PS of the exposed subject (e.g., the difference is less than 0.05). We then fit the outcome regression model among the matched population. In inverse probability treatment weighting, we create a weighted pseudo-population consisting of weighted copies of observed data to remove confounding bias and implement the outcome regression model among the pseudo-population. In all models, we will adjust for site in the outcome regression model to account for across-site heterogeneity. Due to the large number of comparisons, we will select a PS analysis that suits our study setting and data best considering the pros and cons of the different PS-based analyses. We will apply the analysis procedure to each imputed dataset. Suppose $\widehat{\theta_\iota}$ and $\widehat{W_i} \equiv \widehat{var}(\widehat{\theta_\iota})$ denote the point and variance estimates for the parameter of interest θ_0 from the ith imputed dataset, $i = 1,2, \dots, m$. Then the final point estimate $\hat{\theta} = \frac{1}{m}\sum_{i=1}^{m} \widehat{\theta_\iota}$, the average of the m complete-data estimates; and the final variance estimate $\widehat{W} = \frac{1}{m}\sum_{i=1}^{m} \widehat{W_\iota} + \left(1 + \frac{1}{m}\right)\frac{1}{m-1}\sum_{i=1}^{m}\left(\widehat{\theta_\iota} - \hat{\theta}\right)^2$ which accounts for both within- and between-imputation variance.

Handling Pregnancy: Because pregnancy events have an important impact on body weight. We will ignore any weight measurements that occur during a pregnancy period, as identified by ICD-9 and CPT-4 codes. The pregnancy period will be defined as the 9 months before and 3 months after any code indicating a full-term delivery, pre-term delivery, miscarriage, or abortion procedure.

2. Secondary Analysis

As a secondary analysis **[IR-5]**, we will repeat the primary analysis using a distributed-analytic method that only requires sharing of summary-level information. The only difference is we will analyze percent weight changes (from baseline) at 1, 3, and 5 years separately as the analytic tools to fit mixed effects models in distributed databases do not exist yet. Specifically, we will use the distributed linear regression method [5, 6] to compare percent weight changes at 1 year (or 3 or 5 years) between surgery groups. We will consider either PS regression or PS matching as they allow us to use the distributed linear regression approach to conduct analysis using aggregate-level information only. In distributed linear regression, the parameter estimates and covariance matrix have a closed form solution and can be calculated by combining intermediate, summary-level statistics from all sites. [7]
The importance of the secondary analysis is two-fold. First, it will allow us to include additional sites that are not able to share their patient-level data for the primary analysis, thereby allowing us to examine the robustness of our results in a much larger sample size. Second, this will be a key infrastructure activity as it will assess the feasibility and validity of using summary-level information to perform the same analysis that is conventionally done with pooled patient-level information. If proven feasible for PCORnet, the analytic method can further enhance the functionalities of the system by allowing sites to share less sensitive information while preserving the scientific rigor of the study.

There is strong evidence of a temporal trend in the use of the bariatric procedures, and the characteristics of patients who undergo these procedures may also change over time. As a sensitivity analysis **[IR-5]**, we will look at the effects of these procedures on BMI changes by calendar year.

B. AIM 2 (IMPROVEMENTS IN DIABETES RISK)

1. Primary analysis

The analytic framework for Aim 2 is similar to that described in Aim 1. The primary analysis will have three pairwise comparisons. The primary outcome of interest is: (1) remission from diabetes (a time-to-event outcome), defined as HbA1c <6.5% after 90+ days of not having any active anti-hyperglycemic prescription order or dispensing post-surgery. We will estimate the PSs as described in Aim 1, and fit a Cox proportional hazards model to estimate the effects of surgery on diabetes remission, with a maximum follow-up of 1, 3, or 5 years after surgery **[CI-3]**. In addition to regression, matching, and weighting considered in Aim 2, PS stratification is also compatible with Cox regression. Specifically, we can fit a stratified Cox regression stratifying on site and site-specific PS percentiles (e.g., deciles). The secondary outcome of interest is the differential impact of the procedures on glycemic control, measured as change in HbA1c from baseline to 1, 3, and 5 years. As this secondary outcome is a continuous variable, our analytic approach will be similar to the one described in Aim 1.

Handling Pregnancy: Because pregnancy events – particularly the third trimester – can have significant effects on glycemic control, we will ignore any HbA1c measures that occur during a pregnancy period, defined as 9 months before and 3 months after any code indicating a full-term delivery, pre-term delivery, miscarriage, or abortion procedure. We will also ignore any HbA1c measurements that occur during an inpatient hospitalization or during a treatment episode for oral steroid medications. We will also ignore diabetes medication use 9 months before and 3 months after any code indicating a full-term delivery, pre-term delivery, miscarriage, or abortion procedure.

Given the nature of the date ranges for the CDM data, we will not likely be able to construct a meaningful measure of Diabetes Duration at the time of surgery.

C. AIM 3 (MAJOR ADVERSE EVENTS)

1. Primary Analysis

The study cohort in Aim 3 will be similar to the study population in Aim 1. To be eligible for Aim 3, the patient must be linked to state or national death index for mortality outcome or linked to payer data (insurance claims) for other adverse event outcomes. The primary outcome of interest will be subsequent reoperation/reintervention, defined as any additional bariatric procedure and other procedures related to device removals, gastric revisions, abdominal or incisional hernia repair, laparoscopy or laparotomy, and percutaneous endoscopic gastrostomy tube placements. Secondary outcomes of interest are (1) composite end point of 30-day major adverse outcomes, based on the definition used in the LABS study, [12] which includes death; venous thromboembolism; percutaneous, endoscopic, or operative subsequent intervention; and failure to be discharged from the hospital, (2) any hospitalization following initial surgery, and (3) death.

Handling Pregnancy: We will ignore hospitalizations that are related to pregnancy events.

We will explore the use of the total volume of bariatric procedures (or specific bariatric procedures) by site as a potential variable for this adverse events analysis.

Primary analysis: A Cox proportional hazards model will be used to estimate the hazards ratio of each procedure type on time-to-reoperation/re-intervention for three pair-wise comparisons (RYGB vs. SG, RYGB vs. AGB, SG vs. AGB). A single Cox regression model will be estimated with RYGB as the reference procedure. Estimates of the log relative hazard of reoperation for AGB vs. SG will be calculated as the difference in each procedure's log-HR estimate (relative to RYGB) and the delta method will be used to get correct standard errors around that point estimate. The Cox regression model will also include covariate adjustment for key baseline covariates (seeaim3-descriptive-2018-05-28.xlsx for covariates to be included and factor level definitions)). In addition to reporting hazard ratios, we will report the probability of reoperation/re-intervention at 1-, 3-, and 5-years post-surgery with each procedure type for the "average" patient.

We will not adjust for propensity score decile in the outcome regression model (as in Aims 1 and 2) to simplify the analysis and interpretation of results; propensity score adjustment did not influence inference in Aims 1 and 2 and, instead, complicated presentation of the analysis and results. Estimating a single regression model (rather than three pairwise models) makes it easier to explain the estimated risk of an event at 1, 3, and 5 years for an "average" patient, since average is now defined with respect to the entire cohort rather than with respect to the subset of patients with the comparator procedures.

HTE analyses will be performed for the primary reoperation/reintervention outcome; order of HTE analysis listed at the top of analysis plan indicates priority. A sensitivity analysis will also be done where the definition of reoperation/reintervention is expanded to include endoscopy.

As a descriptive analysis, we will also report the unadjusted probability of each constituent type of reoperation and re-intervention (cholecystectomy/ostomy, conversion/revision/reversal, hernia, reoperation, vascular access, and endoscopy) at 1-, 3-, and 5-years post-surgery for each procedure type using estimates from a Kaplan Meier curve.

2. Secondary Analysis

Secondary analysis of time-to-event for all-cause mortality and rehospitalization will follow the same approach as the primary analysis. No HTE or sensitivity analyses will be done.

Secondary analysis of the 30-day composite AE outcome will be done with a logistic regression model. A single logistic regression model will be estimated with RYGB as the reference procedure and SG and AGB as exposures of interest. The log-odds ratio comparing SG to AGB will be calculated as the different in the log-odds ratio comparing each procedure to RYGB, and the delta method will be used to calculate correct standard errors. (Since we only have one outcome per individual, outcomes are assumed independent conditional on covariates and we do not need to use GEE.) The logistic regression model will also be adjusted for baseline covariates and site.

D. AIMS 4-6 (VARIABILITY IN SURGERY OUTCOMES BY BASELINE DEPRESSION, RACE/ETHNICITY, AND AGE)

Aims 4-6 will conduct heterogeneity of treatment (HTE) analyses for primary and secondary outcomes in Aims 1-3. Analytic procedures will be similar to those described above for each outcome's respective aim with a few exceptions noted here. We will not use propensity score adjustment for outcomes in Aims 1 and 2 for the same reasons cited in the analysis plan for Aim 3: that results of our analyses were not affected by propensity score adjustment but the analytic procedures and presentation of results was made more challenging. We will also not repeat sensitivity analyses for Aims 1-3 because they all showed that model findings were robust to the assumptions in question.

For HTE analyses, we will include an interaction between procedure type and the baseline characteristic of interest (depression or race/ethnicity) in the specified regression model. We will conduct ANOVA tests comparing models with and without the interaction to quantify evidence of HTE. Since several outcomes are being examined, we will not have type I error control at $p<0.05$ and will consider analyses to be exploratory. Although HTE analyses for Aim 5 (race and ethnicity) and Aim 6 (Age 65+ vs <65) have already been completed for primary outcomes, we will repeat the analyses in the subset of CDRNs that agree to additional analysis of data and we will add the additional secondary outcomes outlined as above.

HANDLING OF CONFOUNDERS AND MISSING DATA

A. KNOWN CONFOUNDERS AND UNMEASURED POTENTIAL CONFOUNDERS

The clinically rich information available within CDRNs allows us to adjust for many confounders (e.g., BMI, HbA1c) that are often not available in administrative claims databases. To further minimize unmeasured confounding, we require patients to have non-missing key baseline covariate information for age, sex, and baseline BMI **[CI-4]**. We will use PSs to account for these confounders in the analysis. PSs are a proven method that has been shown to provide valid estimates in observational studies.

Update for Aims 4-6: We will not use propensity score methods in Aims 4-6 to simplify analysis and interpretation of results. Use of propensity score adjustment did not affect results in Aims 1-2, but using propensity score methods made analysis and interpretation challenging. Without propensity score adjustment, we can estimate a single regression model that adjusts for procedure arm (rather than estimating three models for three pairwise comparisons). Pooling information across all three procedures increases the power to identify differences between procedures by adding precision to estimated covariate effects. This approach also simplifies interpretation of findings since it does not produce two (possibly different) estimates of outcome rates at time periods of interest (1, 3, and 5 years) from the two separate pairwise comparison models for that procedure.

B. MISSING DATA

We require that covariate information on sex, age, and baseline BMI be present for all patients included in the analysis. Some baseline information, including blood pressure, race, and ethnicity may be missing. For blood pressure, we categorized continuous systolic and diastolic blood pressure into a categorical variable indicating hypertension status and included a "missing". For race and ethnicity, we included

"missing" as a possible value for each variable. Sensitivity analyses for Aims 1-3 demonstrated that findings were not sensitivity to excluding people with missing blood pressure, race, or ethnicity information at baseline.

Other covariate information is defined by the presence of particular diagnosis codes or medications in a patient's medical record and, thus, it is not possible to distinguish between whether a patient is missing particular covariate information (e.g., diabetes diagnosis or insulin prescription) versus a patient who does not have that diagnosis or prescription. There are no missing data methods to adjust for this potential missingness (MD-2). For Aim 1, we conducted a sensitivity analysis that included only data from integrated health systems, where all diagnostic and prescription data are captured through a combination of claims and clinical data, and found that estimated comparisons were similar (MD-4).

Our analyses include a variety of outcome types, including longitudinal, binary, and time-to-event data. Regression models for each make different assumptions about missing outcome information. Mixed effects regression models, used for change in weight (Aim 1) and change in HbA1c (Aim 2) outcomes assume that the observation of the outcome during follow-up is "missing at random" (MAR), that is, that outcome observation is not informative of that outcomes value after conditioning on baseline characteristics and previous observations. In this setting, if a patient is likely to stop getting their weight measured at the onset of a trajectory change (e.g., the patient stops losing weight or begins losing much more weight), then this assumption is violated. There are no missing data methods to address this type of missingness (missing not at random or MNAR) (MD-2). Sensitivity analyses restricted to integrated health systems will include more thorough follow-up since all weight measures will be captured (rather than just weight measurements taken at the health system that performed the bariatric procedure) and, for integrated systems, disenrollment from the health plan would need to be associated with informative missingness for MAR assumptions to be violated (MD-4). Survival analyses with time-to-event data also rely on a MAR assumption: censoring is assumed to be independent of future risk of an event conditional on baseline covariates and observed survival time. Duration of follow-up will be reported for all patients (MD-3).

ADDRESSING HETROGENTITY OF TREATMENT EFFECTS (HTE)

Analyses of HTE will allow us to gain insight into whether the examined associations are consistent across clinically relevant populations (e.g., across key demographic groups). **[HT-1]**. To assess possible HTE, we propose to repeat the primary analyses for all 3 aims within each of the subgroups defined by race/ethnicity (e.g., non-Hispanic black, Hispanic, non-Hispanic white), age (adolescents will be analyzed separately; also <65 years and 65 years), ≥ baseline BMI (≥50 vs <50 kg/m^2), baseline smoking status prior to surgery (current, former, other), and comorbidity groups (e.g., baseline diabetes) or comorbidity index score quartiles **[HT-2, RQ-4].** When there are two subgroups (e.g., BMI ≥50 vs. BMI <50 kg/m^2), we will examine the difference between the subgroup-specific effect estimates (mean difference in Aim 1, and log (hazard ratio) in Aims 2 and 3). The variance of the difference is the sum of the variances for the subgroup-specific effect estimates as the two effect estimates are independent. The standardized difference is expected to follow a normal distribution asymptotically. When there are more than two

subgroups, the differences between the effect estimate from a reference subgroup and the other subgroup-specific effect estimates, after appropriate linear transformations, follows a multivariate normal distribution asymptotically. Then the sum of the differences squared follows a chi-square distribution with the degree of freedom being the number of subgroups minus 1. [21] A significant difference in treatment effect is detected if the corresponding p-value is ≤ 0.05 **[HT-3]**.The purpose of the HTE analysis is to explore any possible treatment effect heterogeneity instead of formal hypothesis testing. Thus we propose to use the same 0.05 significance level for all contrasts **[HT-4]**.

For Aims 4-6, HTE analyses will be repeated for all primary and secondary outcomes in Aims 1-3.

PLAN FOR DATA QUERIES

PCORnet has set up a distributed research model in which a main operating principle is to "send questions to the data." Under this principle, the PCORnet Coordinating Center sends honed programming code ("queries") to each CDRN node site, who run them, unaltered, against the Common Data Model. The study team combines the aggregate-level data from the multiple sites for analysis. Advantages over traditional pooled analysis are that data are kept in the hands of the original data holders, which decreases proprietary concerns and data breaches. Data holders transfer only minimum necessary data to a central site for analysis. In addition, data holders know their data well, facilitating cleaning and troubleshooting. In the proposed study, we will adhere to these principles while at the same time comparing findings from a fully distributed programming approach that yields only aggregate data with an approach that employs (de-identified) individual-level data. We propose a data flow system that takes advantage of, and integrates, the knowledge and expertise of the 3 main components of the study. The data flow system is shown in Figure 1, below.

The scientific teams for the studies will be responsible for defining the analytic plans including defining all outcomes, exposures, and covariates of interests. This information will be passed along to the Methods Core, which will be responsible for creating functional specifications and, working with the DRN OC Query Fulfillment Team, the technical specifications. After multiple iterations, involving the study teams, the Query Fulfillment Team will send the technical specifications to the StatLog team for code development.

Once developed, the Query Fulfillment Team will internally test queries on manufactured data sets, test the query in the pilot datamart, further beta test the query in willing datamarts, and then develop a production-level code package, based on feedback from the beta test sites. When prepared, and with confirmed data use agreements and IRB approvals from sites, the Query Fulfillment Team will transmit the code package using PopMedNet. PopMedNet™ is the query management platform for the PCORnet Distributed Research Network. In PCORnet's distributed data environment, code is developed centrally and distributed to each partner to execute against data that are stored in a common format.

HANDLING SPECIAL FORMS OF OBESITY

It was proposed that we should look (in ICD-9 codes) for but not exclude "special forms" of obesity, such as congenital conditions (e.g., Prader-Willi), especially in the adolescent patient population. We will

identify these patients and adjust for the conditions in the analysis. We will be hesitant to look at this group separately if it is small for practical considerations (i.e., avoiding additional analyses in the setting of limited time and resources). There is a need to develop a list of ICD-9 codes to identify these "special forms" of obesity. Below is a list or genetic disorders commonly associated with hyperphagia and the early development of obesity. The first 5 are typically present as obesity with developmental delay, while the last 5 present with obesity without developmental delay. In addition, craniopharyngioma patients (I.e. S/P surgical resection) should be included.

1. Prader-Willi
2. Albright Hereditary Osteodystrophy
3. Bardet–Biedl Syndrome
4. BDNF and TRKB Deficiency
5. SIM1 Deficiency
6. MC4R Deficiency
7. Leptin and Leptin Receptor Deficiency
8. POMC processing disorders
9. PCSK1 Deficiency
10. SH2B1 Deficiency
11. Craniopharyngioma
12. ROHHAD
13 Alstrom

POTENTIAL ADDITIONAL SECONDARY OR EXPLORATORY ANALYSES:

Time and resources permitting, we will consider conducting a number of additional secondary or exploratory analyses for our aims; however, we currently believe that these analyses are either a) beyond the scope of our current proposal, or b) are likely to be challenged by lack of sufficient data to address the question.

a. It was proposed that we conduct a secondary analysis among patients with a BMI between 30 and 34.9 kg/m^2 because this subgroup is of high interest to the bariatric community, especially as it relates to diabetes outcomes, but also in relation to weight loss and safety.
b. Look at geographical variation across the US for the 3 procedures
c. Look at hospital volume and look for correlation with outcomes and complications.
d. For Aim 3, opportunity to analyze outcomes according to the following designation of centers: low vs. high volume centers, accredited vs non-accredited centers, academic vs community centers, teaching vs non-teaching hospitals.
e. Consider the impact of people who are on weight loss medications after surgery and what differential impact this may have on our weight loss and diabetes outcomes across procedures, including: phentermine, orlistat (alli/Xenical), lorcaserin (Belviq), phentermine HCL/topiramate ER (Qsymia), Naltrexone HCL/bupropion HCL ER (Contrave), and Liraglutide (victoza and Saxenda). We may want to consider Metformin as well.
f. For the adolescent group, it may be interesting to compare surgeries at adolescent centers vs. non-adolescent centers, specifically evaluating surgical complications, weight loss, and co-morbidity improvement.

TABLES AND FIGURES

Table 1. Common Data Model Data Tables that are Applicable to the PCORnet Bariatric Study

PCORnet CDM Data Table	Key Information/Description of Variables	Date Range
DEMOGRAPHIC	Key demographic variables include: Age at surgery (not date of birth), sex, Hispanic ethnicity (y/n) & race are captured.	2004 – 2016* * includes information occurring 1 year before the first bariatric procedure (2005) and one year after the last bariatric procedure (2015) in the study
ENROLLMENT	Enrollment is a concept that defines a period of time during which all medically-attended events are expected to be observed. This concept is often insurance-based, but other methods of defining enrollment are possible. Key variables include: Enrollment start and end dates, and enrollment basis (which relates to how the enrollment period was defined, e.g., insurance, geography, algorithmic, encounter)	2004 – 2016
ENCOUNTER	Contains 1 record for each time a patient sees a provider in ambulatory setting or is hospitalized; multiple encounters per day are possible if they occur with different providers or in different care settings. Encounter type will be used to identify initial bariatric procedures and all subsequent complications and procedures during the follow-up period. Encounters may also be used to determine enrollment periods. Provider, facility information, and DRGs may be used to calculate propensity scores. Key variables include: encounter ID, admit date and time, discharge date and time, provider ID (a pseudoidentifier), facility location (3 digit zip code) and ID, encounter type (ambulatory, emergency, etc), discharge disposition (which can indicate death during hospitalization), and DRG.	2004 – 2016
DIAGNOSIS	Contains all uniquely recorded diagnoses for all encounters. Each diagnosis is associated with a specific patient and encounter. Diagnosis codes and associated encounter dates will be used to establish medical history prior to surgery, to calculate propensity scores, and to identify adverse events. The main focus will be on obesity related diagnoses and adverse events, such as diabetes and hypertension. Key variables include: Diagnosis date,	2004 – 2016

	Diagnosis, diagnosis type (e.g., ICD-9, ICD-10), diagnosis source, and principal diagnosis flag	
PROCEDURES	Procedure codes and associated encounter dates will be used to establish bariatric surgery dates and any reoperations or reinterventions. Key variables include: Procedure date, Procedure, Procedure type (e.g., ICD-9, ICD-10), procedure source	2004 – 2016
VITAL	Contains one record for each recorded height, weight, and blood pressure, which may be included in our propensity scores, as well ask information on smoking status, which is an important comorbid health indicator for the propensity score. Key variables include: height, weight, body mass index (BMI), systolic and diastolic blood pressure, smoking status, tobacco use & type, vital source, and dates of those measures	2004 – 2016
DISPENSING	This data table contains outpatient pharmacy dispensing information. Not all CDRNs and study sites will have access to these data, but many will. These data will be used to capture the use of diabetes treatments, and may be used to identify other comorbid health conditions and/or important co-occurring treatments, such as 2nd generation antipsychotics and obesity drugs, which can have an important impact on body weight. Key variables include: dispensing date, National Drug Code (NDC), days supply, and dispense amount	2004 – 2016
LAB_RESULT_CM	This table contains information on common laboratory measures. For this study, the laboratory measure of interest is the HbA1c for identifying diabetes and diabetes control over time. Key variables include: Lab name, specimen source, specimen date, result, and result unit	2004 – 2016
CONDITION	This data table contains information on a patient's diagnosed and self-reported health conditions and diseases. The patient's medical history and current state may both be represented. This may be used to capture medical comorbidities prior to surgery, including diabetes, as well as complications of surgery. Key variables include: condition, condition status, report date, onset date, resolve date, condition type, condition source	2004 – 2016
PRO_CM	This table contains patient-reported outcome (PRO) common measures which are standardized measures that are defined in the same way across all PCORnet networks. Each measure is recorded at the individual item level: an individual question/statement, paired with its standardized response options. While PRO data are not part of the specific aims of this study, PCORI has required us to report on the presence/absence of PRO data in our	2004 – 2016

	bariatric population to guide possible future studies. Key variables include: PRO item, PRO date, PRO response, PRO method, and PRO mode	
PRESCRIBING	This data table includes information on provider orders for medication dispensing or administration. It will be used to capture the use of diabetes drug treatments, and may be used to identify other comorbid health conditions and/or important co-occurring treatments, such as 2nd generation antipsychotics and obesity drugs, which can have an important impact on body weight. Key variables include: Rx med name, Rx order date, Rx quantity, Rx refills, RxNORM (drug identifier), Rx days supply, Rx frequency, Rx basis (dispensed or administered)	2004 – 2016
DEATH	This table contains reported mortality for patients in the study. This is critical information for Aim 3. Key variables include: death date, death date impute, death source, death match confidence	2004 – 2016
DEATH_CAUSE	This table contains cause of death information for patients in the study. This information will be used for Aim 3. Key variables include: Death cause, code, type, source, and confidence	2004 – 2016
HARVEST	This table identifies attributes associated with each site's specific PCORnet data mart implementation, which will be helpful for data quality checks. Key variables include: data mart ID, name, platform, CDM version, claims use, EHR use, management variables (indicate where imputation or obfuscation were used for variable creation), refresh date (when data were most recently loaded)	NA

Table 2. ICD-9 and CPT-4 codes used to identify bariatric procedures from the CDM

Code	Description	Procedure Assignment	Code type
43.82	Laparoscopic vertical (sleeve) gastrectomy	SG	ICD-9
43.89	Partial gastrectomy with bypass gastrogastrostomy; Sleeve resection of stomach	SG	ICD-9
43775	Laparoscopic sleeve gastrectomy	SG	CPT-4
44.31	High gastric bypass; Printen and Mason gastric bypass	RYGB	ICD-9
44.39	Other gastroenterostomy; Bypass gastroduodenostomy; gastroenterostomy; gastrogastrostomy; Gastrojejunostomy without gastrectomy NOS	RYGB	ICD-9
43633	Gastrectomy, partial, distal; with Roux-en-Y reconstruction	RYGB	CPT-4
43846	Gastric restrictive procedure, with gastric bypass, for morbid obesity; with short limb (less than 100 cm) Roux-en-Y gastroenterostomy	RYGB	CPT-4
43847	Gastric restrictive procedure, with small intestine reconstruction to limit absorption; with long limb (>150 cm) Roux-en-Y	RYGB	CPT-4
44.38	Laparoscopic gastroenterostomy; Bypass: gastroduodenostomy; gastroenterostomy; gastrogastrostomy; Laparoscopic gastrojejunostomy without gastrectomy NEC	RYGB	ICD-9
43644	Laparoscopy, surgical, gastric restrictive procedure with gastric bypass and Roux-en-Y gastroenterostomy (roux limb 150 cm or less)	RYGB	CPT-4
43645	Laparoscopy, surgical, gastric restrictive with gastric bypass and small intestine reconstruction to limit absorption	RYGB	CPT-4
43844	Laparoscopic gastric restrictive procedure with gastric bypass and Roux en Y gastroenterostomy	RYGB	CPT-4
S2085	Lap GASTRIC BYPASS	RYGB	HCPC
43843	Gastric restrictive procedure, without gastric bypass, for morbid obesity; other than vertical-banded gastroplasty	AGB	CPT-4
44.95	Laparoscopic gastric restrictive procedure Adjustable gastric band and port insertion	AGB	ICD-9
43770	Laparoscopy, surgical, gastric restrictive procedure: placement of adjustable gastric band	AGB	CPT-4
S2082	Lap Band	AGB	HCPC

Procedure Assignment Rules: We will assign the specific bariatric procedure type based on ICD-9, CPT-4, and HCPCS procedure codes and the following rules that have been successfully applied in our prior studies:

i. If only one code is present on the same day for the first bariatric procedure for a given patient, we assume that is the correct procedure assignment
ii. If two or more codes are present on the same day for the first bariatric procedure for a given patient and they correspond to the same procedure type (e.g., RYGB, AGB, SG) then we assume that is the correct procedure assignment

Code	Description	Procedure Assignment	Code type
	iii. It two or more codes are present on the same day and they disagree on procedure type, we exclude that case		

Table 3. Revisional and uncommon bariatric procedures to be excluded if recorded as first procedure in the CDM

Name	Code Type	CPT-4 /ICD-9 Code	Description
AGB REMOVAL	CPT-4	43772	LAP RMVL GASTR ADJ DEVICE
AGB REMOVAL	CPT-4	43774	LAP RMVL GASTR ADJ ALL PARTS
AGB REMOVAL	CPT-4	43887	REMOVE GASTRIC PORT, OPEN
AGB REMOVAL	ICD-9	44.97	LAPAROSCOPIC REMOVAL OF GASTRIC RESTRICTIVE DEVICE(S)
AGB REPLACEMENT	CPT-4	43773	LAP REPLACE GASTR ADJ DEVICE
AGB REPLACEMENT	CPT-4	43888	CHANGE GASTRIC PORT, OPEN
AGB REVISION	CPT-4	43771	LAP REVISE GASTR ADJ DEVICE
AGB REVISION	CPT-4	43886	REVISE GASTRIC PORT, OPEN
AGB REVISION	ICD-9	44.96	LAP REVISION OF GASTRIC RESTRICTIVE PROCEDURE
REVISION NOS	CPT-4	43848	REVISION GASTROPLASTY
REVISION NOS	CPT-4	43860	REV GASTROJEJ ANASTOM; W/O VAGOT
VBG	CPT-4	43842	GASTROPLASTY FOR OBESITY
VBG	ICD-9	44.68	LAPAROSCOPIC GASTROPLASTY
BPD	CPT-4	43845	GASTRIC RESTRICTIVE PROCEDURE WITH PARTIAL GASTRECTOMY, PYLORUS-PRESERVING DUODENOILEOSTOMY AND ILEOLIEOSTOMY (BPD)

Table 4. Diabetes Medications List

Insulin Medications	Non-insulin Diabetes Medications
INSULIN ADMINISTRATION SU	ACARBOSE
INSULIN DETEMIR	ACETOHEXAMIDE
INSULIN GLARGINE	CHLORPROPAMIDE
INSULIN HUMAN, RDNA ORIGIN	EXENATIDE
INSULIN ISOPHANE (HUMAN)	GLIMEPIRIDE
INSULIN ISOPHANE HUMAN	GLIMEPIRIDE/PIOGLITAZONE
INSULIN ISOPHANE PORK PURE	GLIPIZIDE
INSULIN LISPRO	GLIPIZIDE/METFORMIN
INSULIN LISPRO/INSULIN, PROTAMINE LISPRO	GLYBURIDE
INSULIN NPH S-S/REG INSULN S-S	GLYBURIDE/METFORMIN
INSULIN REGULAR HUMAN	LINAGLIPTIN
INSULIN ZINC HUMAN	LIRAGLUTIDE
INSULIN, ASPART PROTAMINE, HUMAN/INSULIN	METFORMIN
INSULIN, ASPART, HUMAN	METFORMIN HCL
INSULIN, GLULISINE, HUMAN	METFORMIN HYDROCHLORIDE
INSULIN, PROMPT ZINC, BEEF-PORK	METFORMIN/PIOGLITAZONE
INSULIN, REGULAR, BEEF-PORK	METFORMIN/ROSIGLITAZONE
INSULIN, REGULAR, PORK	METFORMIN/SITAGLIPTIN
INSULIN, ZINC, HUMAN	MIGLITOL
LENTE INSULIN, BEEF	NATEGLINIDE
LENTE INSULIN, BEEF-PORK	PIOGLITAZONE
NPH INSULIN, BEEF	PIOGLITAZONE HCL
NPH INSULIN, BEEF-PORK	PIOGLITAZONE HCL-GLIMEPIR
NPH INSULIN, HUMAN	PRAMLINTIDE
HUMAN/REGULAR INSULIN, HUMALOG	REPAGLINIDE
NPH INSULIN, PORK	ROSIGLITAZONE
NPH, HUMAN INSULIN ISOPHANE/INSULIN REGU	SAXAGLIPTIN
REGULAR INSULIN, HUMAN	SITAGLIPTIN
ULTRALENTE INSULIN, BEEF	TOLAZAMIDE
ULTRALENTE INSULIN, BEEF-PORK	TOLINASE
ULTRALENTE INSULIN, HUMAN	TROGLITAZONE
INSULIN DEGLUDEC,	ALBIGLUTIDE
INSULIN GARGINE U-300	DULAGLUTIDE
U-500 REGULAR HUMAN INSULIN	CANAGLIFLOZIN
	DAPAGLIFLOZIN
	EMPAGLIFLOZIN

Table 5. PCORnet Bariatric Study Outcomes

Aim 1	Aim 2	Aim 3
Primary Outcome: Percentage change in kg in adults (BMI z-score change in adolescents) at 1, 3, and 5 years Secondary Outcomes: 1) weight regain* at 3 and 5 years; estimated as percent regain from the maximum weight (in kg) loss in the first 2 years; 2) a post-operative body weight that is <5% lower than the pre-surgical weight at 1, 3, 5 years after bariatric surgery; and 3) proportion achieving >5%, >10%, >20%, and >30% weight loss at 1, 3, and 5 years	Primary Outcomes: 1) Rate of diabetes remission (HbA1c <6.5% off diabetes medications); Secondary Outcomes: 1) Rate of diabetes relapse after initial remission (restart of medication or HbA1c ≥6.5% 2) Change in HbA1c at 1, 3, and 5 years	1. Primary Outcomes: 1) reoperation/ re-intervention is defined as any additional bariatric procedure and other procedures related to device removals, gastric revisions, abdominal or incisional hernia repair, laparoscopy or laparotomy, and percutaneous endoscopic gastronomy tube placements. We will look at endoscopy as a sensitivity analysis. *Follow-up time is defined as the number of days after surgery until the first reoperation or re-intervention procedure code (if observed) or censoring.* Secondary Outcomes: 1. All-cause mortality is defined as any death during the study period. We will also examine the cause of death. *Follow-up time is defined as the number of days after surgery until death (if observed) or censoring.* 2. 30-day composite adverse event is defined using the LABS study and includes death; venous thromboembolism; percutaneous, endoscopic, or operative subsequent intervention; and, failure to be discharged from the hospitalization where the bariatric surgery was performed (within 30-days of surgery).

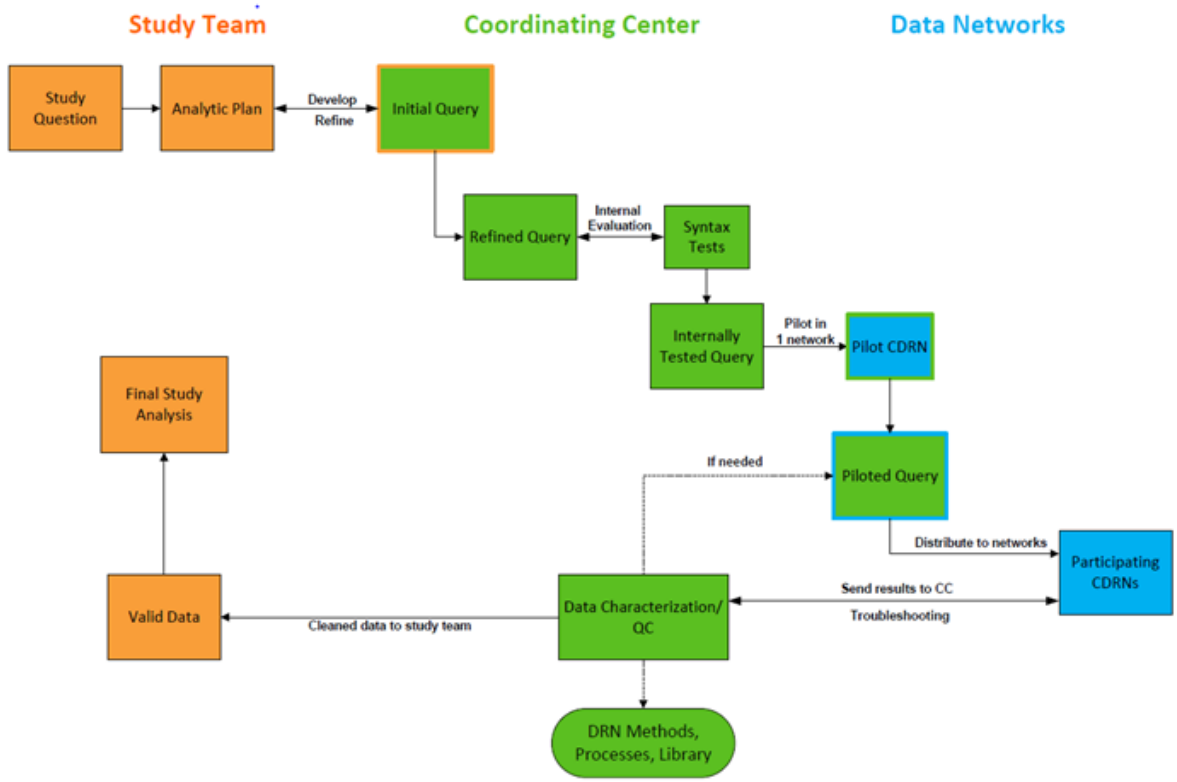
		3. Rehospitalization is defined as any inpatient hospitalization following surgery that is not associated with a delivery, miscarriage, or abortion procedure code (using our previously defined list). *Follow-up time is defined as the number of days after surgery until rehospitalization (if observed) or censoring.*

Table 6. Important Baseline Covariates that will be Considered as Potential Confounders

Variable	Source (CDM Table)	Rationale
Age*	v1.0 DEMOGRAPHIC table; Field Name = BIRTH_DATE	Choice of other surgical procedures vary with age; surgical weight loss and surgical complications likewise vary by age
Baseline BMI*	v1.0 VITAL table; Field Names = HT, WT (date of measurement: Field Name = MEASURE_DATE for each measure)	Initial BMI may shape patient and providers' procedure choice; surgical weight loss and impact on diabetes varies with anthropometric measures.
HbA1c	V2.0/3.0 LAB table; Field Name = A1C (date of measurement: Field Name = RESULT_DATE for each measure)	For Aim 2, HbA1c as a measure of glycemic control at baseline is strongly associated with risk of remission and relapse of diabetes
Diabetes Medication Use	V2.0/3.0 PRESCRIBING (and DISPENSING) table; Field Name = NDC	For Aim 2, insulin use preoperatively is strongly associated with risk of remission and relapse of diabetes For measuring use of Diabetes Medications at baseline, we will create **three variables**: Insulin use Y/N, Oral medication use Y/N, and number of oral medications used
Hospital length of stay 1 year before surgery	V1.0 ENCOUNTER table; Field Names = ADMIT_DATE, DISCHARGE_DATE	This measure is another surrogate for health status, which strongly impacts procedure choice and is associated with surgical outcomes
Comorbid conditions*	v1.0 DIAGNOSIS table; Field Name = DX v1.0 CONDITION table; Field Name = CONDITION	Health status strongly impacts procedure choice and is associated with surgical weight loss and complications of surgery
Sex*	v1.0 DEMOGRAPHIC table; Field Name = SEX	Choice of surgical procedure may vary with sex, and male sex is associated with adverse outcomes
Race/ethnicity*	v1.0 DEMOGRAPHIC table; Field Name = RACE or HISPANIC	Medical/surgical treatment choices often vary with race/ethnicity as does surgical weight loss and surgery's impact on diabetes
Pre-op smoking status*	v1.0 VITAL table; Field Name = TOBACCO	Smoking status may influence procedure choice and impacts diverse health outcomes
CDRN/node site	Meta-data (CDRN of origin)	Regional eating and physical activity norms are likely to be associated with CDRN and with surgical outcomes

** hypothesized effect moderators*

Figure 1. PCORnet Bariatric Study Data Flow System



REFERENCES

1. Inge TH, Zeller MH, Jenkins TM, et al. Perioperative outcomes of adolescents undergoing bariatric surgery: the Teen-Longitudinal Assessment of Bariatric Surgery (Teen-LABS) study. *JAMA pediatrics.* 2014;168(1):47-53.
2. Walker AM, Patrick AR, Lauer MS, et al. A tool for assessing the feasibility of comparative effectiveness research. *Journal of comparative effectiveness research.* 2013;3:11-20.
3. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol.* 2011;64(7):749-759.
4. Sturmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. *Am J Epidemiol.* 2010;172(7):843-854.
5. Karr AF, Lin X, Sanil AP, Reiter JP. Secure regression on distributed databases. *J Comput Graph Stat.* 2005;14(2):263-279.
6. Karr AF, Fulp WJ, Vera F, Young SS, Lin X, Reiter JP. Secure, privacy-preserving analysis of distributed databases. *Technometrics.* 2007;49(3):335-345.
7. Zhang F, Li L, Abrams AM, Kleinman K. Development and implementation of secure statistical analyses on distributed databases. 2010; http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=486. Accessed March 17, 2014.
8. Fireman B, Lee J, Lewis N, Bembom O, van der Laan M, Baxter R. Influenza vaccination and mortality: differentiating vaccine effects from bias. *Am J Epidemiol.* 2009;170(5):650-656.
9. Toh S, Reichman ME, Houstoun M, et al. Multivariable confounding adjustment in distributed data networks without sharing of patient-level data. *Pharmacoepidemiol Drug Saf.* 2013;22(11):1171-1177.
10. Toh S, Reichman ME, Houstoun M, et al. Comparative risk for angioedema associated with the use of drugs that target the renin-angiotensin-aldosterone system. *Arch Intern Med.* 2012;172(20):1582-1589.
11. Toh S, Shetterly S, Powers JD, Arterburn D. Privacy-preserving Analytic Methods for Multisite Comparative Effectiveness and Patient-centered Outcomes Research. *Med Care.* 2014;52(7):664-668.
12. 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.
13. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf.* 2006;15(5):291-303.
14. Lash TL, Fox MP, Fink AK. *Applied Quantitative Bias Analysis in Epidemiologic Data.* Springer; 2009.
15. Rubin DB. Inference and missing data. *Biometrika.* 1976;63(3):581-592.
16. Rubin DB. *Multiple imputation for nonresponse in surveys.* New York, NY: John Wiley & Sons; 1987.
17. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med.* 1991;10(4):585-598.
18. Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. . *Journal of the American Statistical Association.* 1994;89:447-482.
19. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health.* 2006;60(7):578-586.

20. Toh S, Garcia Rodriguez LA, Hernan MA. Analyzing partially missing confounder information in comparative effectiveness and safety research of therapeutics. *Pharmacoepidemiol Drug Saf.* 2012;21 Suppl 2:13-20.
21. Casella G, Berger RL. *Statistical Inference.* Pacific Grove, CA: Duxbury Press; 2001.