Supplementary Online Content 2

Aminian A, Zajichek A, Arterburn DE, et al. Association of metabolic surgery with major adverse cardiovascular outcomes in patients with type 2 diabetes and obesity. *JAMA*. doi:10.1001/jama.2019.14231

STUDY PROTOCOL
(January 6, 2019)

- **Objective**
  To assess the association of bariatric and metabolic surgery in patients with type 2 diabetes mellitus (T2DM) with various individual and composite cardiovascular outcomes in a matched-cohort analysis

- **Data source**
  The data source for all analysis will be Cleveland Clinic's Electronic Health Record (EHR) through 12/31/2018.

- **Study protocol**

I. Cohort derivation

All patients at the Cleveland Clinic who have (or had) T2DM will be considered for study entry. This will serve as the *base cohort*, though all patients initially considered may not be included in the final statistical analysis (see **II.** for description of the algorithm for identifying T2DM). Starting with the *base cohort*, the final study cohort will be derived in two (2) phases:

(1) Identifying bariatric surgeries

All patients in the *base cohort* receiving bariatric surgery who meet the following specifications at the time of surgery will be **included:**

a) Any of the following codes for bariatric procedures:
   - CPT: 43633, 43634, 43770, 43775, 43644, 43645, 43659, 43842, 43843, 43844, 43845, 43846, 43847
   - ICD: 44.31, 43.82, 44.95, 43.89, 44.38, 44.39, 44.68
   - HCPCS: S2082, S2085

   (Chart review will be performed to verify the bariatric and metabolic procedures)
b) Has T2DM (HbA1c > 6.4% *or* on diabetes medication)
c) Age between 18-80 years
d) BMI $\geq$ 30 kg/m^2
   - The closest BMI measurement *prior* will be used. If no measurements are available, the patient will be excluded.
e) Occurred through 12/31/2017
f) Surgery at one (1) of the following Cleveland Clinic locations
   - Ohio (Main Campus, Fairview, Hillcrest, Akron), or Florida (Weston) facilities

After inclusion, patients who meet any of the following at the time of surgery will be **excluded:**

a) History of liver, heart or lung transplant
b) Emergency department admission within five (5) days *prior*
c) Active cancer/GI cancer diagnosis code within one (1) year *prior*
d) Surgical procedures/diagnoses codes for upper GI cancer and/or peptic ulcer disease during the hospitalization for bariatric surgery
e) History of severe heart failure
   - Any ejection fraction <20% *prior*

The date of the *first* bariatric surgery meeting inclusion/exclusion will be identified for each patient and serve as the ***index date***.

(2) Identifying non-surgical comparators

Based on the ***index dates*** derived in **(1)**, similar patients who did not receive bariatric surgery will be obtained for comparison. Starting with the *base cohort*, the following algorithm will be implemented

a) Remove patients identified in **(1)**
   - Patients who met inclusion but were subsequently excluded will also be removed

b) Randomly assign each patient a single date from the collection of ***index dates***. This will serve as the ***assigned date*** for non-surgical patients.
c) **Exclude** patients meeting any of the following at the ***assigned date***:
    - Last follow-up on or before the assigned date
    - Already dead or died within 30 days *after* the assigned date
        - Serves as a proxy for potential terminally-ill patients
    - Does not have T2DM (HbA1c < 6.5% *and* not on diabetes medication)
    - Age < 18 or Age > 80 years
    - BMI < 30
        - The closest measurement *prior* will be used. If none are available, the closest measurement in absolute time will be used.
    - History of liver, heart or lung transplant
    - ED admission within 5 days *prior*
    - Active cancer/GI cancer diagnosis code within one (1) year *prior*
    - History of severe heart failure
        - Any ejection fraction <20% *prior*

The patients identified in **(1)** and **(2)** will form the final study cohort in which subsequent analysis will be performed. The algorithm described in **(2)** will be replicated 5 times for a sensitivity analysis to assess the variability of baseline characteristics of the non-surgical patients at the assigned date, as well as overall model effect estimates and prediction accuracy (See blow, Sensitivity Analyses). At this point, the ***index date*** and ***assigned date*** will collectively be referred to as the ***index date***.

(3) Matching Process

Each non-surgical patient will be matched with a propensity-score ('logit' link) to five (5) surgical patients based on the index date, age at index date, BMI at index date (categorized as 30-35, 35-40, 40+), gender, insulin use, location (Ohio versus Florida), diabetes complications (composite of coronary artery disease, heart failure, cerebrovascular disease, peripheral vascular disease, neuropathy, nephropathy, or being on dialysis) with the MatchIt package in R.[1]

II. Identifying patients with T2DM

A T2DM prevalence cohort has been previously created using the modified Kho algorithm[2] and with the following specifications:

Based upon ICD9/10 codes, patients with type 1 diabetes mellitus (T1DM) and T2DM were identified. Codes for ketoacidosis were classified as T1DM code.

The algorithm was used to calculate the earliest date when a patient record contained any of the following combinations:

- T2DM code in the EHR & T2DM medication
- T2DM code in the EHR & abnormal glucose*****
- T2DM medication & abnormal glucose*
- T2DM code reordered twice in the EHR & outpatient insulin
- Insulin preceded by T2DM medication

*Abnormal glucose: Random glucose >200 mg/dL or fasting glucose >125 mg/dl.
The earliest date that any of the five (5) conditions above were met was documented as the date in which the patient first met the criteria for T2DM in the Cleveland Clinic system.

This cohort contains approximately **N = 287438** patients whose T2DM dates range from 1/1/1998 through 12/31/2017.

III. Outcomes

Patients will be followed starting at the ***index date*** for 6 outcomes as follows (see eTable 1 for diagnosis codes/definitions):

1. All-cause mortality
   - If deaths are unknown (i.e. not entered into Cleveland Clinic's EHR), the *Social Security* and *State Death Indices* will be searched.

2. Coronary artery disease
   - Myocardial infarction
     - ICD9: 410.X, 410.XX
     - ICD10: i21.X and i21.XX, i22.X and i22.XX, i23.X and i23.XX
   - Other
     - ICD9: 411.X, 411.X AND 414.X, 36.01, 36.02, 36.03, 36.05, 36.06, 36.07, 36.10, 36.11, 36.12, 36.13, 36.14, 36.15, 36.16, 36.17, 36.19, 36.31, 36.32, 36.33, 36.64
     - CPT: 92982, 92984, 92995, 92996, 92980, 92981, 33510, 33511, 33512, 33513, 33514, 33516, 33517, 33518, 33519, 33521, 33522, 33523, 33530, 33533, 33534, 33535, 33536, 93539, 93540

3. Cerebrovascular disease
   - Ischemic stroke
     - ICD9: 433.X1, 434.X1
     - ICD10: i63.X and i63.XX and i63.XXX
   - Other
     - ICD9: 436.0, 430.X, 431.X, 38.12, 0.61, 0.63
     - CPT: 37215, 37216, 0075T, 0076T, 35301, 37205, 37206

4. Heart failure
   - ICD9: 428.0, 428.1, 428.20, 428.21, 428.22, 428.23, 428.30, 428.31, 428.32, 428.33, 428.40, 428.41, 428.42, 428.43, 428.9

5. Nephropathy
   - Two (2) measures of eGFR < 60 separated by at least 90 days with no intervening values >= 60
   - eGFR will be computed using the Modification of Diet in Renal Disease (MDRD) Study equation with creatinine levels

6. Atrial fibrillation
   - ICD9: 427.31
   - ICD10: i48.0, i48.2, i48.91
   - CPT: 93650, 93653, 93656, 93657

Primary endpoint

The primary outcome will consist of the composite of 1-6, recording the *first* occurrence after the index date as the event date. The conditions/events in which a patient had in their history at baseline will be omitted from the composite definition in follow-up.

Secondary endpoints

The secondary composite outcome will consist of the composite of (1) all-cause mortality, (2) myocardial infarction, and (3) ischemic stroke, recording the *first* occurrence after the index date as the event date. The conditions/events in which a patient had in their history at baseline will be omitted from the composite definition in follow-up.

Each of 1-6 will also assessed individually. For each outcome, any patient with a history of that outcome at the index date will be eliminated from the risk set. For nephropathy, patients with a baseline eGFR < 60 and patients with history of dialysis will also be excluded.

IV. Data elements

The following data elements will be extracted from the EHR for each patient in the *base cohort* and used for analysis. Collected variables will be mapped to the Unified Medical Language System (UMLS) identifiers.[3] (Note: Specific diagnosis, visit, encounter, procedure, prescription, and lab dates will be collected with each data element, where applicable, to track changes over time)

- Demographics
    - Date of birth
    - Gender
    - Race
    - Median zip code income
    - Weight
    - BMI
    - Insurance
    - Smoking status
    - Location
- Medical history
    - Hypertension
    - Dyslipidemia
    - Coronary artery disease
    - Cerebrovascular events
    - Heart failure
    - Nephropathy
    - Diabetic neuropathy
    - Peripheral arterial disease
    - Atrial fibrillation
    - COPD
    - Being on dialysis
- Labs
    - HbA1c
    - Creatinine
    - eGFR
    - Systolic/diastolic BP
    - LDL
    - HDL
    - Triglyceride levels
    - Urine albumin: creatinine ratio (UACR)
- Medication
    - Insulin
    - Non-insulin diabetes medication
    - ACE-I/ARB
    - Other antihypertensive medication
    - Statin (or other cholesterol-lowering medication)
    - Aspirin
    - Warfarin

V. Statistical analysis

Cause-specific event rates per 100 patient-years of follow-up starting from the index date will be estimated for each outcome within each study group. Cumulative incidence estimates and hazard ratios (HRs) from a fully-adjusted Cox-PH model will be obtained for each outcome. All baseline variables will be included to adjust for potential confounding. The proportional-hazards assumptions for the treatment variable will be tested based on weighted residuals as proposed by Grambsch and Therneau.[4]

Within each outcome dataset, missing values will be multiply-imputed with Multiple Imputation by Chained equations to create 5 imputed datasets. Predictive mean matching, logistic regression, and polytomous logistic regression will be used for numeric, binary, and categorical variables, respectively. Imputation-corrected standard errors of model estimates and contrasts will be obtained by Rubin's formula.[5,6]

The percent (%) of weight loss and HbA1c values will also be evaluated over time and compared among the surgical and non-surgical groups.

A significance level of 0.05 will be considered statistically significant, and 95% CIs will be reported where applicable. Adjustment for multiple comparisons of secondary endpoints is not performed and analyses will be considered as exploratory. All analysis will be done in the R statistical programming language (version 3.5.0).[7]

VI. Sensitivity analyses

Sensitivity of the HR estimates from the fully-adjusted Cox models to two components of the process used in obtaining non-surgical controls will be assessed: random-sampling of index dates, and matching ratio. The sampling of index dates for non-surgical patients will be repeated five (5) times, and matching ratios of 1:1, 1:5, and 1:10 will be used in each resulting dataset. All analyses will be performed on all 15 datasets to assess sensitivity, though results are reported from one (1) dataset using 1:5 matching.

Furthermore, time-varying HRs for surgical versus non-surgical patients at 2, 5, and 8 years following the index date will be estimated. The details of sensitivity analyses and results have been mentioned below.

VII. References

[1] Ho DE, Imai K, King G, Stuart EA. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software. 2011;42(8): 1 - 28.

[2] Kho AN, Hayes MG, Rasmussen-Torvik L, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. J Am Med Inform Assoc. 2012;19(2):212-8.

[3] Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. Ann Transl Med. 2018;6(3):42.

[4] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994;81(3): 515-526.

[5] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 2011; 45(3): 1 - 67.

[6] Frank E Harrell Jr (2018). rms: Regression Modeling Strategies. R package version 5.1-2. https://CRAN.R-project.org/package=rms

[7] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

AMENDMENT TO STUDY PROTOCOL
(August 2, 2019)

In order to address the comments raised during the journal pee-review process, the following parts will be added to the study protocol.

- **Adverse Events of Metabolic Surgery**
  Early (<90 days after surgery) and late adverse events (need for nutritional, endoscopic, radiologic, and surgical interventions until last follow-up to serve as proxy to adverse events) after metabolic surgery will be collected and reported (see eTable 2 for diagnosis and intervention codes/definitions). Nutritional variables of interests are hemoglobin, serum protein, albumin, and vitamin D levels.

- **E-value for Sensitivity Analysis**
  Additional sensitivity analysis to assess the potential effects of unmeasured confounders will be added to the study by the E-value methodology. The E-value is the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and outcome, conditional on the measured covariates, to fully explain away a specific treatment–outcome association.[1-3]
  For all 8 composite and individual study endpoints, the E-value will be calculated and interpreted for the HR estimates and for their limit of 95% CI closest to the null.

- **References**
  [1] VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167(4):268-274.

  [2] Haneuse S, VanderWeele TJ, Arterburn D. Using the E-Value to Assess the Potential Effect of Unmeasured Confounding in Observational Studies. JAMA. 2019;321(6):602-603.

  [3] https://www.hsph.harvard.edu/tyler-vanderweele/tools-and-tutorials/