**Methods:**

*Study population:*

We performed a case- control study approved by the Mayo Institutional Board Review (IRB: 17-003675 including consecutive patients with prevalent EAC from 2009 -2012. We excluded incident (history of BE >1 year before EAC diagnosis) and Siewert III EAC. The BE group was composed of surveillance patients who did not progress to EAC, from the Rochester Epidemiology Project (REP) database years 1991 to 2018 12, 13. Using the International Classification of Disease codes, 9th (ICD-9-CM 530.85) and 10th (ICD-10-CM K22.7).  and confirmed by reviewing the endoscopic and histologic reports. BE was defined as endoscopic evidence of at least 1 cm of visible columnar mucosa in the esophagus and histologic evidence of intestinal metaplasia.

*Data abstraction*

Baseline characteristics, EAC stage, metabolic labs and risk factors were manually extracted from medical records. Metabolic diseases using ICD-9 and ICD-10 codes were confirmed using manual record review. Chronic reflux was defined as the presence of heartburn for more than 6 months as reported by the patient during their clinical visit. Metabolic labs were extracted from the time period closer to the BE and EAC diagnosis without a time limitation. A mean of multiple values was recorded. If ALT elevation due to liver metastases or chemotherapy, this was not used. Nonalcoholic steatohepatitis (NASH) was defined by high BMI (BMI>25) orfatty appearing liver on cross sectional imaging and elevated ALT>40 IU/L with ALT>AST, without alternative explanation. Metabolic syndrome was defined as the presence of three or more of the following risk factors: central obesity, hypertension, elevated LDL, and elevated HgA1c or diabetes 14. BMI ≥30 kg/m2 was considered a surrogate for central obesity 15.

*Statistical analysis:*

The study outcomes were to compare risk factors between BE and all EAC and early EAC (stage I or II) We performed one-to-one matching between EAC and BE based on age and sex to assure comparability between the two groups. using the nearest neighbor method, within a caliper width of 0.1 of the standard deviation of the logit of the propensity score. Baseline characteristics were compared using chi-square χ2 analysis for categorical variables and Student’s t-test for continuous variables. If missing data were more than 5%, we used inverse probability weight to account for missing data. All statistical analyses were performed using STATA 14.2 (College Station, TX: StataCorp LP.)