Evaluation of treatment effect modification in propensity score matching: An empirical example

M. Sanni Ali[1,2]; Ian J. Douglas[1]; Elizabeth Williamson[1]; Daniel Prieto-Alhambra[2,3]; Liam Smeeth[1]

1London School of Hygiene and Tropical Medicine, London, UK;

2 University of Oxford, Oxford, UK; 3 Universitat Autonoma de Barcelona and Instituto de Salud Carlos III, Barcelona, Spain

Background

Propensity score matching (PSM) is a common approach to estimate treatment effects in observational studies. Of interest to in studies is whether treatment effect is modified by some covariate. This is often assessed by stratifying the PSM data by levels of the effect modifier which may break the match resulting in biased treatment effect in strata of the modifier.

Objectives

To assess the impact of stratifying on effect modifier after PSM on balance of covariates and treatment effects

Methods

We used the United Kingdom Clinical Practice Record Datalink (CPRD) data to evaluate the association between bariatric surgery (BS) on the risk of type 2 diabetes mellitus (T2DM) as an example. We followed patients from index date until the earliest of T2DM, death, loss-to follow-up, or end of study (December 31, 2014). We considered cardiovascular disease status (CVD status) as effect modifier of bariatric surgery on T2DM. We applied 1:1 PSM with caliper of 0.20 standard deviation on the logit of the PS. We evaluated covariate balance using absolute standardized mean difference (ASMD) in the full PSM data as well as PSM data stratified by CVD status. We estimated hazard ratios and 95% confidence intervals (95% CI) using Cox proportional hazard models.We also fitted traditional Cox models and compare the treatment effects.

Results

After PSM, we had 5132 patients with 406 having T2DM events during follow-up.When stratified by CVD status, the number of subjects and events in the two groups were 4946/378 and 180/28 with and without CVD, respectively. PSM improved covariate balance, ASMD all below 10%. On the otherhand, stratification after PSM by CVD status broke the match resulting in imbalance in the covariates with ASMD up to 42%. Bariatric surgery reduced the risk of T2DM, HR [95% CIs], 0.65 [0.53, 0.79] in the full matched cohort. Similar risk was detected in those without CVD at baseline HR = 0.65 [0.53, 0.79] but not in patients with CVD, HR = 0.80 [0.37, 1.7]. Traditional Cox models resulted in similar effect estimates, HRs = 0.56 [0.47, 0.66], 0.51[0.47, 0.65], and 0.80 [0.40, 1.52], in the full data, in those with and without CVD, respectively.

Conclusions

Our findings highlight consequence of stratification on levels of effect modifier on covariate balance after PSM. Covariate balance achieved on the full PS-matched data does not transfer to balance within strata of a modifier unless further matching is considered, and we recommend against using this approach for evaluating treatment effect modification in PSM.