Real-World Evaluation of Systematic Bias and Balance of Overall Patient Characteristics of Propensity Score Matching Versus Cardinality Matching

Stephen P Fortin, Martijn J Schuemie

Background

Propensity-score matching (PSM) is susceptible to substantial bias and model overparameterization due to limited overlap in covariate distributions and degrees of freedom, respectively, especially at small sample sizes[1,2,3,4]. Cardinality matching (CM), which finds the largest matched sample meeting a set of prespecified balance criterion, overcomes these limitations by balancing covariates directly[1]. Prior research has shown large-scale CM achieves superior patient retention and comparable systematic bias as compared to large-scale PSM; however, large-scale methods may not be applicable in the setting of small sample sizes[5]. The current study leverages OHDSI tools (OMOP common data model [CDM] and HADES) to compare the performance of PSM versus CM at progressively smaller sample sizes with increasingly limited covariate overlap.

Methods

We identified new users of β-blockers versus angiotensin-converting enzyme inhibitors (ACEI) monotherapy between 10-01-2014 and 01-01-2017 from the IBM® MarketScan® Commercial Claims and Encounters database (index = first drug exposure). PSM was performed through nearest-neighbor matching (1:1, caliper=0.15). CM was performed at four prespecified matching criterion (1:1; max SMD=0.00, 0.01, 0.05 and 0.10 between matching covariates). Matching covariates (n=38), defined as covariates used to fit the propensity score model and perform cardinality matching, included patient demographics and clinical characteristics, comprised of subcomponents of the Charlson comorbidity index. We assessed post-match patient retention; matching covariate balance; and balance on overall patient characteristics, defined by a larger covariate set the majority of which were not included among matching covariates. The larger covariate set was comprised of patient demographics and all observed conditions, drug exposures, procedures and other health-service-use behaviors. Expected absolute systematic error (EASE), a summary measure of expected systematic bias, was assessed using 105 negative control outcome experiments. We created a 10% and 0.25% sample group consisting of 5 and 200 subsample draws, respectively. For each sample group, negative control outcome experiments were conducted across the pooled post-match subsamples.

Results

A total of 186,233 (β-blocker: 56,871; ACEI: 129,362) patients meeting the study criteria were identified; and 18,576 (β-blocker: 5,675; ACEI: 12,901) and 465 (β-blocker: 142; ACEI: 323) patients were included in each subsample draw of the 10% and 0.25% sample groups, respectively. As compared to PSM, CM achieved higher average post-match sample size at all prespecified balance criteria in the 10% sample group. Similarly, in the 0.25% sample group, CM achieved improved post-match patient retention except at the tightest prespecified balance criterion (PSM vs. CM, max SMD=0.00 vs. CM, max SMD=0.10: 232 vs. 178 vs. 269). As shown in Figure 1, CM achieved balance across all matching covariates; while PSM was associated with an average of 4.2 (11.1%) and 7.6 (19.9%) imbalanced matching covariates in the 10% and 0.25% sample groups, respectively.

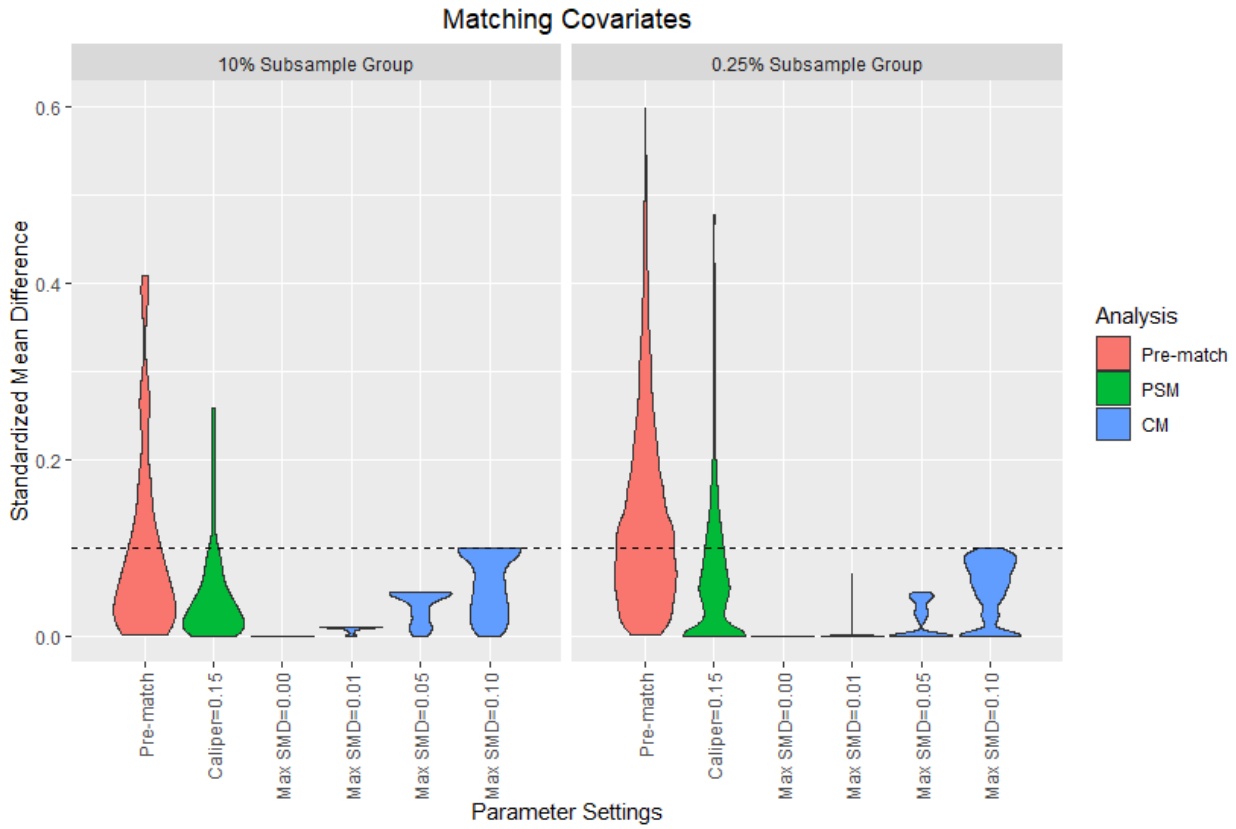

Figure 1. Post-match standardized mean difference of matching covariates with propensity score matching (PSM) versus cardinality matching (CM)

An average of 35,458 and 8,566 covariates were observed and included in the larger covariate set in the 10% and 0.25% sample groups, respectively. In the 10% sample group, fewer observed covariates were imbalanced post-match on average with PSM (2.2%, n=780) as compared to CM (max SMD=0.0: 2.5%, n=896; and max SMD=0.10: 3.1%, n=1,115). On the other hand, in the 0.25% sample group, the average number of post-match imbalanced observed covariate with PSM (16.2%, n= 1,389) was either similar or worse as compared to CM (max SMD=0.0: 13.8%, n=1,178; and max SMD=0.10: 16.9%, n=1,445). The post-match balance among the larger covariate set with PSM versus CM is summarized in Figure 2.

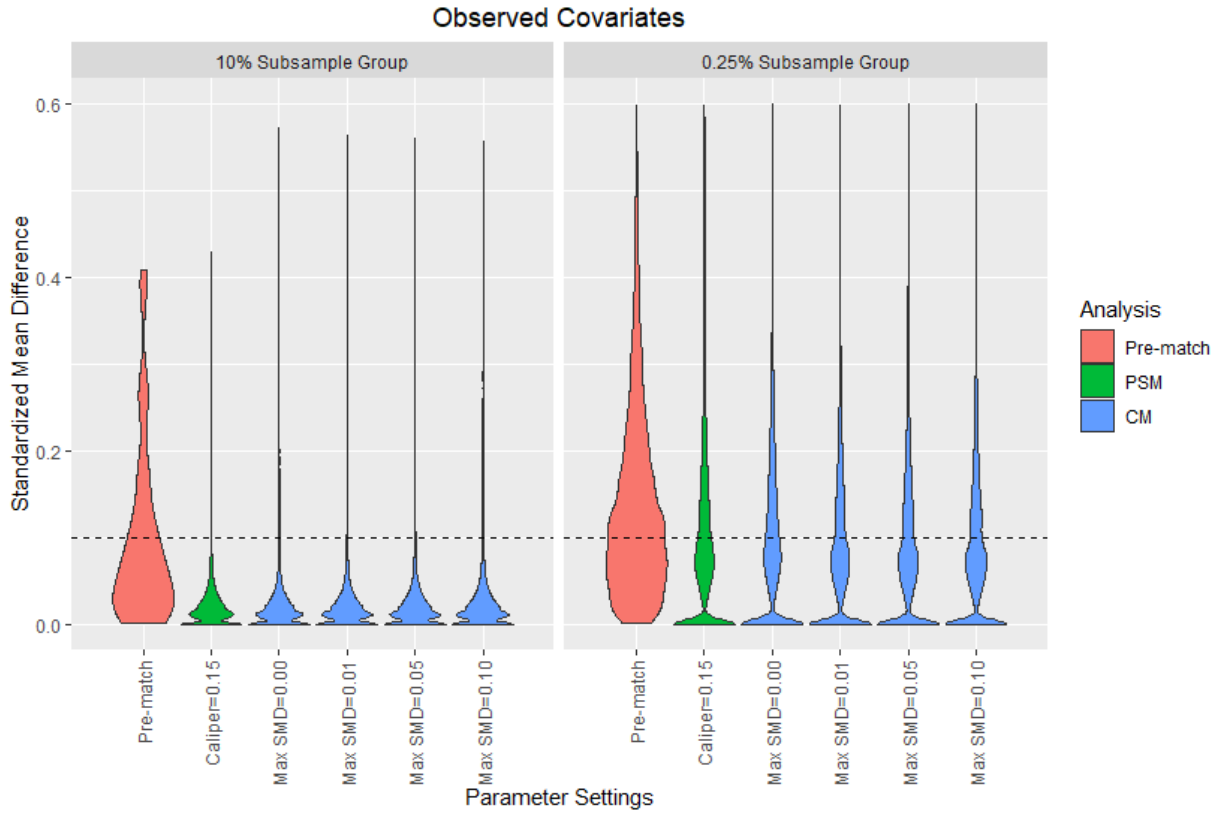

Figure 2. Post-match standardized mean difference of larger covariate set with propensity score matching (PSM) versus cardinality matching (CM)

As shown in Figure 3, both PSM and CM were associated with significant reductions in systematic bias in the 10% and 0.25% sample groups. Comparable reductions in EASE were observed between matching methods in the 0.25% sample group; however, PSM was associated with a greater reduction in EASE in the 10% sample group. Furthermore, tighter prespecified balance criterion were associated with improved reduction in EASE with CM.

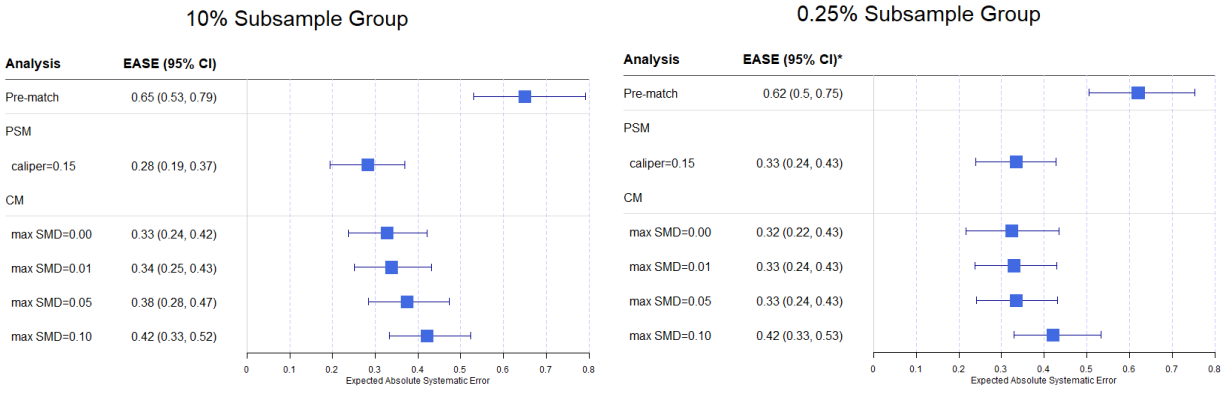

Figure 3. Post-match expected absolute systematic error with propensity score matching (PSM) versus cardinality matching (CM)

Conclusion

CM found the largest matched sample meeting a set of prespecified balance criteria. At smaller sample sizes, PSM and CM achieved comparable balance in overall patient characteristics and reductions in systematic bias albeit CM had improved performance at more stringent prespecified balance criteria (i.e., SMD <0.05). Improved indirect covariate balance and reductions in EASE were observed with PSM at larger sample sizes as compared to CM. We recommend CM as an alternative to PSM in studies of small sample size.

References/Citations

1. Visconti G, Zubizarreta, JR. Handling limited overlap in observational studies with cardinality matching. *Observational Studies*. 2018; 4, 217-249.
2. Crump RK, Hotz VJ, Imbens GW, Mitnik OA. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*. 2009; 96 (1), pp. 187–199.
3. Rothe C. Robust confidence intervals for average treatment effects under limited overlap. *Econometrica*. 2017; 85: 645-660. doi: 10.3982/ECTA13141.
4. Pirracchio R, Resche-Rigon M, Chevret S. Evaluation of the propensity score methods for estimating marginal odds ratios in case of small sample size. *BMC Med Res Methodol.* 2012;12:70. doi:10.1186/1471-2288-12-70.
5. Fortin SP, Johnston SS, Schuemie MJ. Comparison of Cardinality Matching and Propensity Score Matching for Causal Inference in Observational Research. *BMC Med Res Methodol*. 10.21203/rs.3.rs-94412/v1.