

Matching on treatment in observational research – what is the role of the matching algorithm?

Ulrike Held and Priska Heinz

Department of Biostatistics, University of Zurich

March 2021

Abstract

Propensity score matching (PSM) for balancing groups of patients with different treatments has been used extensively in the medical literature in recent years. PSM can be applied to observational data, but also to data from randomized controlled trials (RCTs) and is intended to obtain a causal treatment effect estimate [1]. Current discussions in the literature address the potential presence of unmeasured confounding, double adjustment after matching [2], and the estimand itself, but there is only little information on the properties of the matching algorithm and its effect on the estimated treatment effect. In recent years, different matching algorithms have been proposed, and a systematic evaluation of their properties was lacking. We conducted a simulation study to evaluate the role of the matching algorithm in different scenarios of treatment effect size for clinical applications. The corresponding protocol can be found under <https://osf.io/unbka/>. The results of our simulation study show that full matching can in general be fully supported, genetic and caliper matching can be recommended as well. In this context, we discussed coverage, bias, computing time, and potential loss of power.

Finally, we applied the different matching algorithms to a large multi-center prospective cohort study on lumbar spinal stenosis patients in Switzerland [3]. We estimated the treatment effect of non-operative treatment versus surgical treatment on clinically relevant reduction of pain between baseline and 12-month follow-up by using propensity score matching for treatment decision including 14 relevant confounders reported in the literature.

References

- [1] E. A. Stuart. Matching methods for causal inference: A review and a look forward. *Stat Sci*, 25(1):1–21, Feb 2010.

- [2] P. C. Austin. Double propensity-score adjustment: A solution to design bias or bias due to incomplete matching. *Stat Methods Med Res*, 26(1):201–222, Feb 2017.
- [3] U. Held, J. Steurer, G. Pichierri, M. M. Wertli, M. Farshad, F. Brunner, R. Guggenberger, F. Porchet, T. F. Fekete, U. D. Schmid, I. Gravestock, and J. M. Burgstaller. What is the treatment effect of surgery compared with nonoperative treatment in patients with lumbar spinal stenosis at 1-year follow-up? *J Neurosurg Spine*, pages 1–9, 04 2019.