Observational Studies 4 (2018) 292-296 Submitted 06/18; Published 12/18

Mitigating Reporting Bias in Observational Studies Using Covariate Balancing Methods

Guy Cafri **Guy.X.Cafri@kp.org**
Surgical Outcomes and Analysis
Kaiser Permanente
8954 Rio San Diego Drive, Suite 406, San Diego, CA 92108

Elizabeth W. Paxton **Liz.W.Paxton@kp.org**
Surgical Outcomes and Analysis
Kaiser Permanente
8954 Rio San Diego Drive, Suite 406, San Diego, CA 92108

Keywords: Reporting Bias, Study Modification Bias, Protocol, Blinding, Propensity Score, Covariate Balancing

1. Introduction

Observational studies are commonplace in medical research. Despite their widespread use, one criticism is that they are at increased risk for selective publishing of results, also known as reporting bias. Reporting bias can involve selective publication of a whole study (publication bias) or selective publication of outcomes (outcome reporting bias)(Chan et al., 2004a) (Chan et al., 2004b). A subtler form of reporting bias is selective modification to a study's design and/or analysis, which we refer to as study modification bias. Concerns about reporting bias are not limited to observational studies, they arise in randomized clinical trials as well (Chan et al., 2004a) (Chan et al., 2004b).However, widely adopted guidelines related to conduct and reporting (ICH, 1998) (Schulz et al., 2010), study protocol and statistical analysis plan development (Chan et al., 2013) (Gamble et al., 2017) and detailed information provided on trial registries is thought to limit reporting bias in randomized clinical trials. Parallel efforts have been initiated for observational studies, development of reporting standards (Von Elm et al., 2007), proposed use of statistical analysis plans (Thomas and Peterson, 2012) and registration of observational studies (Loder et al., 2010). Parallel efforts have been initiated for observational studies, development of reporting standards (Von Elm et al., 2007), proposed use of statistical analysis plans (Thomas and Peterson, 2012) and registration of observational studies (Loder et al., 2010). Even with adherence to these recommendations, observational studies remain at risk for study modification bias. For example, in a multivariable regression modeling setting different model specifications, variable and data inclusions (e.g., variable transformations, interactions, exclusion of outliers) might be explored, with the most favorable result likely to be reported (Rubin, 2007). By comparison, first balancing the treatment and comparison groups on covariates and subsequently looking at the relationship with the outcome is less susceptible to study modification bias because the design and analysis phases of the study are kept separate (Rubin, 2007). While covariate balancing methods have the potential to mitigate study modification bias, they too can be exploited. For example, including different subsets of observations and covariates or choosing among different balancing methods until the most favorable treatment-outcome relationship is found. In this communication we describe optimal ways matching and weighting covariate balancing methods might be used to mitigate study modification bias.

2. Developing a Study Protocol

Often in studies of randomized medical interventions all aspects of the design are detailed in a study protocol/statistical analysis plan. Key information to include in an observational study protocol consist of which patients will be included, which covariates will be used and what methods will be applied to attain balance. Two different approaches may be taken to developing a protocol when using covariate balancing methods. In the first approach a protocol is developed prior to covariate balancing. How distance metrics are estimated (e.g., propensity score, Mahalanobis distance) and how they are applied to balance the data (e.g., nearest neighbor matching, optimal matching, weighting by the odds) should be reported. A comparison of approaches using different combinations of distance estimation and application is recommended to improve the prospects of attaining good balance (Harder et al., 2010). If multiple approaches are used and the one that maximizes balance is selected, then balance maximization needs to be operationalized in the protocol. One advantage of this approach is that what is planned is completely independent of the data and therefore may not be prone to reporting bias. However, in some research contexts it may be difficult to provide adequate detail in a protocol without first examining treatment-covariate relations. For example, it may not be possible to attain reasonable balance without excluding observations (e.g., trimming or subsetting; Dehejia and Wahba (1999); Rosenbaum (2012)). If a protocol is developed prior to covariate balancing and observations are ultimately excluded to improve balance, then information pertaining to which patients are included in the study will be absent or inaccurate. One solution to this problem would be to amend the protocol to incorporate updated information, if necessary, after covariate balancing.

In the second approach, covariate balancing methods are applied first and thereafter a protocol is written detailing the methods that were undertaken to balance the data (e.g., Deshpande et al., 2017). The advantages of this approach are a more complete initial protocol and not requiring the protocol authors to anticipate all the problems that may arise in the balancing process. There is at least one important disadvantage associated with this approach. It is possible that after having observed imbalance on one or more covariates, the authors decide to drop those covariates. We recommend that a protocol include all covariates considered prior to balancing and all covariates after balancing was complete and provide a reason for removal of any covariates. There are several important practical issues underpinning the above recommendations. First, we frequently refer to balance as being "good" or "adequate" without explicitly defining what that means. Perhaps the most common rule of thumb for adequate balance is that no covariate should have an absolute standardized difference > 0.10, although a more liberal cut off value of 0.25 has also been applied (Harder et al., 2010). Neither threshold can guarantee the complete removal of bias, particularly if the covariate has a strong relationship with the response (Harder et al., 2010)

(Ho et al., 2007). One way of addressing this issue in the context of matching is prioritizing balance on certain covariates (Pimentel, 2016) suspected of having the strongest relationship to the response. Another way of protecting against bias is to reduce imbalances as much as possible regardless of whether some cutoff has been achieved. Improved balance often results through exclusion of some observations. However, we view this as an option to pursue after all others have been exhausted because of the bias that might ensue (Rosenbaum and Rubin, 1985). In some cases "adequate" balance cannot be achieved no matter what approach is taken. In this context the results of covariate balancing methods should be reported but with the acknowledgement that residual imbalances on measured covariates may be a limitation of the study.

3. Creating a Physical Blind

Use of a physical blind can act as an important safeguard when implementing covariate balancing methods (Yue, 2012). In this approach the outcome is removed from the dataset containing treatment and covariate information but allowing for that information to be linked back to each observation's outcome. In the most conservative approach the person balancing the data never has access to the outcome and is different from the person analyzing the treatment-outcome relationship. A more liberal implementation that would be less resource intensive is having the same individual balance and model the treatment-outcome relationship, but access is granted to the outcome data by an intermediary only after balancing the data.

4. Discussion

In this commentary we have described ways in which reporting bias can be mitigated using covariate balancing methods. The recommendations provided in this communication are not without controversy. Concerns have been raised with protocol-driven observational studies which by association apply to the current recommendations. One argument often levied is that protocols restrict the ability of observational studies to make new discoveries (Sørensen and Rothman, 2010). A counter to this claim is that not all observational studies are exploratory, there is a distinct class of confirmatory observational studies to which these standards do apply (Loder et al., 2010). Moreover, there is nothing to preclude the reporting of a data-driven result in a confirmatory observational study if it is disclosed as such. There are important reasons for considering covariate balancing methods beyond their potential for mitigating reporting bias. The parallels to the design and analysis of a randomized trial are particularly compelling, for example, with respect to the types of inferences made and the magnitude of the estimated treatment effects (Austin, 2014). Implementing covariate balancing methods, in conjunction with recommended reporting guidelines, a study protocol/ statistical analysis plan and study registration can limit reporting bias and improve the quality and transparency of medical studies.

References

Austin, P. C. (2014). The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Statistics in Medicine*, 33(7):1242–1258.

Chan, A.-W., Hróbjartsson, A., Haahr, M. T., Gøtzsche, P. C., and Altman, D. G. (2004a). Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*, 291(20):2457–2465.

Chan, A.-W., Krleža-Jerić, K., Schmid, I., and Altman, D. G. (2004b). Outcome reporting bias in randomized trials funded by the canadian institutes of health research. *Canadian Medical Association Journal*, 171(7):735–740.

Chan, A.-W., Tetzlaff, J. M., Gøtzsche, P. C., Altman, D. G., Mann, H., Berlin, J. A., Dickersin, K., Hróbjartsson, A., Schulz, K. F., Parulekar, W. R., et al. (2013). Spirit 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ*, 346:e7586.

Dehejia, R. H. and Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *JAMA*, 94(448):1053–1062.

Deshpande, S. K., Hasegawa, R. B., Rabinowitz, A. R., Whyte, J., Roan, C. L., Tabatabaei, A., Baiocchi, M., Karlawish, J. H., Master, C. L., and Small, D. S. (2017). Association of playing high school football with cognition and mental health later in life. *JAMA neurology*, 74(8):909–918.

Gamble, C., Krishan, A., Stocken, D., Lewis, S., Juszczak, E., Doré, C., Williamson, P. R., Altman, D. G., Montgomery, A., Lim, P., et al. (2017). Guidelines for the content of statistical analysis plans in clinical trials. *JAMA*, 318(23):2337–2343.

Harder, V. S., Stuart, E. A., and Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods*, 15(3):234.

Ho, D. E., Imai, K., King, G., and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15(3):199–236.

ICH (1998). *ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials E9.* European Medicines Agency, London, England.

Loder, E., Groves, T., and MacAuley, D. (2010). Registration of observational studies. *BMJ*, 340(2):c950.

Pimentel, S. D. (2016). Large, sparse optimal matching with r package rcbalance. *Observational Studies*, 2:4–23.

Rosenbaum, P. R. (2012). Optimal matching of an optimally chosen subset in observational studies. *Journal of Computational and Graphical Statistics*, 21(1):57–71.

Rosenbaum, P. R. and Rubin, D. B. (1985). The bias due to incomplete matching. *Biometrics*, pages 103–116.

Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Statistics in Medicine*, 26(1):20–36.

Schulz, K. F., Altman, D. G., and Moher, D. (2010). Consort 2010 statement: updated guidelines for reporting parallel group randomized trials. *Annals of Internal Medicine*, 152(11):726–732.

Sørensen, H. T. and Rothman, K. J. (2010). The prognosis for research. *BMJ*, 340.

Thomas, L. and Peterson, E. D. (2012). The value of statistical analysis plans in observational research: defining high-quality research from the start. *JAMA*, 308(8):773–774.

Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., Initiative, S., et al. (2007). The strengthening the reporting of observational studies in epidemiology (strobe) statement: guidelines for reporting observational studies. *PLoS Medicine*, 4(10):e296.

Yue, L. Q. (2012). Regulatory considerations in the design of comparative observational studies using propensity scores. *Journal of Biopharmaceutical Statistics*, 22(6):1272–1279.