Duke | Robert J. Margolis, MD Center for Health Policy

# Understanding the Use of Negative Controls to Assess the Validity of Non-Interventional Studies of Treatment Using Real-World Evidence

**Virtual Public Workshop**
**March 8, 2023, 10:00 AM – 3:00 PM ET**

## Overview of Key Negative Control Techniques

| Method | Brief Description | Key Assumptions* | Strengths | Limitations |
|---|---|---|---|---|
| **Bias detection/adjustment via NCE**[1-4, 29,30] | In a regression model of outcome on treatment, NCE, and measured covariates, the presence of an association between NCE and outcome implies residual confounding, while a null association implies no empirical evidence of residual confounding. Under certain assumptions, coefficient of NCE equals the unmeasured confounding bias. | - Linear additive outcome model<br><br>- The association between NCE and unmeasured confounder is equal to the association between treatment and unmeasured confounder | - Intuitive and easy to implement in practice | - Strong modeling assumptions<br><br>- Only leverage one type of negative controls |
| **Bias detection/adjustment via NCO**[3-5, 27,28] | In a regression model of NCO on treatment and measured covariates, the presence of an association between NCO and treatment implies residual confounding, while a null association implies no empirical evidence of residual confounding. Under certain assumptions, coefficient of treatment in the NCO model equals the unmeasured confounding bias. NCO has also been used for bias adjustment in survival analysis. | - Linear additive outcome model<br><br>- The association between NCO and unmeasured confounder is equal to the association between outcome and unmeasured confounder | - Intuitive and easy to implement in practice<br><br>- Connects to traditional difference-in-differences method | - Strong modeling assumptions<br><br>- Only leverage one type of negative controls |

| | | | | |
|---|---|---|---|---|
| **Large-scale NC pairs for detection or calibration[6-10]** | By estimating the effect of exposure on outcomes across a collection of settings where the exposure is not believed to cause the outcome, one can estimate an empirical null distribution of the exposure effect and compute calibrated p-values that take both random and systematic error into account. | - Bias follows a normal distribution whose mean and variance can be corrected estimated using negative drug-outcome pairs<br><br>- Distribution assumption only holds for calibration | - Intuitive and easy to implement in practice<br><br>- Utilizes the rich drug-outcome information in EHR data | - Strong distributional assumption<br><br>- Validation of the large number of negative drug-outcome pairs selected |
| **Control outcome calibration (COCA) using NCO[11,32]** | Search for the causal effect (constant additive effect[11] or nonparametric identification of the average treatment effect on the treated[32]) such that the NCO-treatment association is null, adjusting for covariates and Y(0). | - Enriching the adjustment set of covariates with the potential outcome under no treatment, Y(0), suffices to adjust for confounding between NCO and treatment | - Leverages the NCO to search for the right amount of treatment effect | - Relies on the conditional independence assumption<br><br>- Only leverage one type of negative controls |
| **(Generalized) difference-in-differences using NCO[12-13]** | The difference-in-difference method adjusts for unmeasured confounding leveraging the baseline outcome which is an NCO. There is also a scale-invariant generalization of the difference-in-differences method. | - Confounding of NCO-treatment relationship equals the confounding of outcomes-treatment on the quantile scale | - Leverages the baseline outcome which is widely available as NCO to adjust for confounding bias | - Relies on additional model assumption<br><br>- Only leverage one type of negative controls |
| **Double negative control method[14-26,31]** | Also referred to as proximal causal learning in the literature. Leverage an NCO and an NCE to identify causal effect subject to unmeasured confounding without any modeling restriction. Methods have been developed for point exposure[14,15,17], discrete setting[16], longitudinal setting[15,18], survival analysis[19], mediation analysis[20], panel data setting[21,22,31], heterogeneous treatment effect[23], dynamic treatment regime[24], test-negative | - NCO and NCE provide sufficient information about the unmeasured confounder | - Leverages a pair of NCs to fully identify bias; no modeling assumption required, allows for flexible modeling, provides double robustness methods, and | - Need to identify an NCO and an NCE |

| | | | | |
|---|---|---|---|---|
| | design[25], outcome-dependent sampling[26]. | | applies to a range of settings | |
| **Data-driven automated negative control estimation (DANCE)[33]** | Search for triplets of disconnected NCs then aggregate all candidate NC pairs to estimate the average treatment effect | - Linear structural equation model<br><br>- Disconnected NCs: NCs causally related to neither the treatment nor the outcome | - Data-driven selection and validation of negative control<br><br>- Estimates causal effect combining all NC pairs | - Strong model assumption |

* Only listing key assumptions in addition to the assumption that the selected NCE and/or NCO variables are valid NC = negative control; NCE = negative control exposure; NCO = negative control outcome
** This overview table of key negative control techniques was developed by Dr. Xu Shi.

**Bibliography**


1. Flanders WD, Klein M, Darrow LA, Strickland MJ, Sarnat SE, Sarnat JA, Waller LA, Winquist A, Tolbert PE. A method for detection of residual confounding in time-series and other observational studies. Epidemiology (Cambridge, Mass.). 2011 Jan;22(1):59.
Smith GD, Lipsitch M, Tchetgen ET, Cohen T. Negative control exposures in epidemiologic studies. Epidemiology. 2012 Mar 1;23(2):350-2.

2. Weisskopf MG, Tchetgen Tchetgen EJ, Raz R. Commentary: on the use of imperfect negative control exposures in epidemiologic studies. Epidemiology. 2016 May 1;27(3):365-7.

3. Arnold BF, Ercumen A, Benjamin-Chung J, Colford Jr JM. Brief report: negative controls to detect selection bias and measurement bias in epidemiologic studies. Epidemiology (Cambridge, Mass.). 2016 Sep;27(5):637.

4. Shi X, Miao W, Tchetgen Tchetgen EJ. A selective review of negative control methods in epidemiology. Current epidemiology reports. 2020 Dec;7:190-202.

5. Sanderson E, Richardson TG, Hemani G, Davey Smith G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. International journal of epidemiology. 2021 Aug;50(4):1350-61.

6. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. Statistics in medicine. 2014 Jan 30;33(2):209-18.

7. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Robust empirical calibration of p-values using observational data. Statistics in medicine. 2016 Sep 9;35(22):3883.

8. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. Proceedings of the National Academy of Sciences. 2018 Mar 13;115(11):2571-7.

9. Schuemie MJ, Ryan PB, Hripcsak G, Madigan D, Suchard MA. Improving reproducibility by using high-throughput observational studies with empirical calibration. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2018 Sep 13;376(2128):20170356.

10. Gruber S, Tchetgen Tchetgen EJ. Limitations of empirical calibration of p-values using observational data. Statistics in medicine. 2016 Sep 30;35(22):3869-82.

11. Tchetgen Tchetgen EJ. The control outcome calibration approach for causal inference with unobserved confounding. American journal of epidemiology. 2014 Mar 1;179(5):633-40.

12. Glynn A, Ichino N. Generalized nonlinear difference-in-difference-in-differences. V-Dem Working Paper. 2019 Jun 1;90.

13. Sofer T, Richardson DB, Colicino E, Schwartz J, Tchetgen Tchetgen EJ. On negative outcome control of unobserved confounding as a generalization of difference-in-differences. Statistical science: a review journal of the Institute of Mathematical Statistics. 2016;31(3):348.

14. Miao W, Geng Z, Tchetgen Tchetgen EJ. Identifying causal effects with proxy variables of an unmeasured confounder. Biometrika. 2018 Dec 1;105(4):987-93.

15. Tchetgen Tchetgen EJ, Ying A, Cui Y, Shi X, Miao W. An introduction to proximal causal learning. arXiv preprint arXiv:2009.10982. 2020 Sep 23.

16. Shi X, Miao W, Nelson JC, Tchetgen Tchetgen EJ. Multiply robust causal inference with double-negative control adjustment for categorical unmeasured confounding. Journal of the Royal Statistical Society. Series B, Statistical methodology. 2020 Apr;82(2):521.

17. Cui Y, Pu H, Shi X, Miao W, Tchetgen Tchetgen EJ. Semiparametric proximal causal inference. arXiv preprint arXiv:2011.08411. 2020 Nov 17.

18. Ying A, Miao W, Shi X, Tchetgen Tchetgen EJ. Proximal causal inference for complex longitudinal studies. arXiv preprint arXiv:2109.07030. 2021 Sep 15.

19. Ying A, Cui Y, Tchetgen Tchetgen EJ. Proximal causal inference for marginal counterfactual survival curves. arXiv preprint arXiv:2204.13144. 2022 Apr 27.

20. Dukes O, Shpitser I, Tchetgen Tchetgen EJ. Proximal mediation analysis. arXiv preprint arXiv:2109.11904. 2021 Sep 24.

21. Tchetgen Tchetgen EJ, Dukes O, Shi X, Miao W, Richardson D. RE:"SYNTHETIC CONTROL METHODS FOR THE EVALUATION OF SINGLE-UNIT INTERVENTIONS IN EPIDEMIOLOGY: A TUTORIAL". American journal of epidemiology. 2022 May;191(5):965-6.

22. Shi X, Miao W, Hu M, Tchetgen Tchetgen EJ. Theory for identification and inference with synthetic controls: a proximal causal inference framework. arXiv preprint arXiv:2108.13935. 2021 Aug 31.
23. Sverdrup E, Cui Y. Proximal Causal Learning of Heterogeneous Treatment Effects. arXiv preprint arXiv:2301.10913. 2023 Jan 26.

24. Qi Z, Miao R, Zhang X. Proximal learning for individualized treatment regimes under unmeasured confounding. Journal of the American Statistical Association. 2022 Nov 14(just-accepted):1-33.

25. Li KQ, Shi X, Miao W, Tchetgen Tchetgen EJ. Double negative control inference in test-negative design studies of vaccine effectiveness. ArXiv. 2022 Mar 23.

26. Li KQ, Shi X, Miao W, Tchetgen Tchetgen EJ. Doubly Robust Proximal Causal Inference under Confounded Outcome-Dependent Sampling. arXiv preprint arXiv:2208.01237.

2022 Aug 2.

27. Richardson DB, Laurier D, Schubauer-Berigan MK, Tchetgen ET, Cole SR. Assessment and indirect adjustment for confounding by smoking in cohort studies using relative hazards models. American journal of epidemiology. 2014 Nov 1;180(9):933-40.

28. Tchetgen Tchetgen EJ, Sofer T, Richardson D. Negative outcome control for unobserved confounding under a Cox proportional hazards model. Available at https://biostats.bepress.com/harvardbiostat/paper192/.

29. Flanders WD, Strickland MJ, Klein M. A new method for partial correction of residual confounding in time-series and other observational studies. American journal of epidemiology. 2017 May 15;185(10):941-9.

30. Miao W, Tchetgen Tchetgen EJ. Invited commentary: bias attenuation and identification of causal effects with multiple negative controls. American journal of epidemiology. 2017 May 15;185(10):950-3.

31. Imbens G, Kallus N, Mao X. Controlling for unmeasured confounding in panel data using minimal bridge functions: From two-way fixed effects to factor models. arXiv preprint arXiv:2108.03849. 2021 Aug 9.

32. Tchetgen Tchetgen EJ, Park C, Richardson D. Single Proxy Control. arXiv preprint arXiv:2302.06054. 2023 Feb 13.

33. Kummerfeld E, Lim J, Shi X. Data-driven Automated Negative Control Estimation (DANCE): Search for, Validation of, and Causal Inference with Negative Controls. arXiv preprint arXiv:2210.00528. 2022 Oct 2.

*This public workshop is supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) as part of a financial assistance award [U19FD006602] totaling $4,241,714 with 100 percent funded by FDA/HHS. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by FDA/HHS, or the U.S. Government.*