

Biblio Test

Emma Tarmey

February 2024

## **1 Introduction**

## References

- [1] Jason Abrevaya, Yu-Chin Hsu, and Robert P. Lieli. Estimating Conditional Average Treatment Effects. *Journal of Business and Economic Statistics*, 33(4):485–505, 2015. Publisher: American Statistical Association.
- [2] Alexandre Belloni and Victor Chernozhukov and Christian Hansen. High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives*, 28(2):29–50, 2014.
- [3] Anirban Basu. Estimating Person-Centered Treatment (PeT) Effects Using Instrumental Variables: An Application to Evaluating Prostate Cancer Treatments. *Journal of Applied Econometrics*, 29(4):671–691, 2014. Publisher: John Wiley & Sons, Ltd.
- [4] Bica, Ioana and Alaa, Ahmed, M and Jordon, James and van der Schaar, Mihaela. Estimating Counterfactual Treatment Outcomes over Time Through Adversarially Balanced Representations. *In Proc. 8th International Conference on Learning Representations (ICLR 2020)*, 2020.
- [5] Howard D. Bondell and Brian J. Reich. Simultaneous Regression Shrinkage, Variable Selection, and Supervised Clustering of Predictors with OSCAR. *Biometrics*, 64(1):115–123, 2008. Publisher: International Biometric Society.
- [6] Leo Breiman. Random Forests. *University of California*, 2001. Publisher: University of California, Berkeley CA, 94720.
- [7] Leo Breiman. Statistical Modelling: Two Cultures. *Statistical Science*, 16(3):199–231, 2001.
- [8] Gregor Buch, Andreas Schulz, Irene Schmidtman, Konstantin Strauch, and Philipp S. Wild. A systematic review and evaluation of statistical methods for group variable selection. *Statistics in Medicine*, 42(3):331–352, 2022. Publisher: John Wiley & Sons Ltd.
- [9] Gregor et al. Buch. A systematic review and evaluation of statistical methods for group variable selection. *Statistics in Medicine*, 42:331–352, 2023.
- [10] Andrea Burton, Douglas G. Altman, Patrick Royston, and Roger L. Holder. The design of simulation studies in medical statistics. *Statistics in Medicine*, 25, 2006. Publisher: Wiley InterScience.
- [11] Chen, Hua Yun and Li, Hesen and Argos, Maria and Persky, Victoria W. and Turyk, Mary E. Statistical Methods for Assessing the Explained Variation of a Health Outcome by a Mixture of Exposures. *International Journal of Environmental Research and Public Health*, 19(5):2693, 2022.

- [12] Hugh A. Chipman, Edward I. George, and Robert E. McCulloch. BART: Bayesian Additive Regression Trees. *The Annals of Applied Statistics*, 4(1):266–298, 2010. Publisher: Institute of Mathematical Statistics.
- [13] Mohammad Ziaul Islam Chowdhury and Tanvir C Turin. Variable selection strategies and its importance in clinical prediction modelling. *Family Medicine and Community Health*, 8(1), 2020. Publisher: BMJ Specialist Journals \_eprint: <https://fmch.bmj.com/content/8/1/e000262.full.pdf>.
- [14] Carlos Cinelli, Andrew Forney, and Judea Pearl. A Crash Course in Good and Bad Controls. *Journal of Sociological Methods and Research*, 2022.
- [15] Diego Colombo, Marloes H. Maathuis, Markus Kalisch, and Thmoas S. Richardson. Learning High-Dimensional Directed Acyclic Graphs with Latent and Selection Variables. *The Annals of Statistics*, 40(1):294–321, 2012. Publisher: Institute of Mathematical Sciences.
- [16] Daniel, Rhian and Zhang, Jingjing and Farewell, Daniel. Making apples from oranges: Comparing noncollapsible effect estimators and their standard errors after adjustment for different covariate sets. *Biometrical Journal*, 63(3):528–557, 2020.
- [17] T. Daniya, M. Geetha, B. Santhosh Kumar, and R. Cristin. Least Squares Estimation of Parameters for Linear Regression. *International Journal of Control and Automation*, 13(2):447–452, 2020. Publisher: SERSC.
- [18] de Luna, Xavier and Fowler, Philip and Johansson, Per. Proxy Variables and Nonparametric Identification of Causal Effects. *Institute for the Study of Labor*, 10057, July 2016.
- [19] Edward P.K. Parker and Elsie M.F. Horne and William J. Hulme and John Tazare and Bang Zheng and Edward J. Carr and Fiona Loud and Susan Lyon and Viyaasan Mahalingasivam and Brian MacKenna and Amir Mehrkar and Miranda Scanlon and Shalini Santhakumaran and Retha Steenkamp and Ben Goldacre and Jonathan A.C. Sterne and Dorothea Nitsch and Laurie A. Tomlinson. Comparative effectiveness of two- and three-dose COVID-19 vaccination schedules involving AZD1222 and BNT162b2 in people with kidney disease: a linked OpenSAFELY and UK Renal Registry cohort study. *The Lancet Regional Health*, 30, July 2023.
- [20] Efron, Bradley and Hastie, Trevor and Jognstone, Iain and Tibshirani, Robert. Least Angle Regression. *The Annals of Statistics*, 32(2):407–499, April 2004.
- [21] Eklund, Martin and Zwanzig, Silvelyn. SimSel: a new simulation method for variable selection. *Journal of Statistical Computation and Simulation*, 82(4):515–527, 2012.

- [22] Elizabeth J. Williamson, Zoe Aitken, Jock Lawrie, Shyamali C. Dharmage, John A. Burgess, Andrew B. Forbes. Introduction to causal diagrams for confounder selection. *Respirology*, 19(3):303–311, 2014.
- [23] T. Emmanuel, T. Maupong, and D. et al Mpoeleng. A survey on missing data in machine learning. *Journal of Big Data*, 8(140), 2021.
- [24] Pressat-Laffouilhère et al. Variable selection methods were poorly reported but rarely misused in major medical journals: Literature review. *Journal of Clinical Epidemiology*, 139:12–19, 2021.
- [25] F. Richard Guo and Qingyuan Zhao. Confounder Selection via Iterative Graph Expansion. *arXiv*, October 2023.
- [26] Fan, Jianqing and Li, Runze. Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association*, 96(456):1348–1360, 2001.
- [27] J. Max Findlay, Joshua Nisar, and Tim Darsaut. Cerebral Vasospasm: A Review. *Canadian Journal of Neurological Sciences*, 43:15–32, 2016. Publisher: Cambridge University Press.
- [28] C.L. Francoeur and S.A. Mayer. Management of delayed cerebral ischemia after subarachnoid hemorrhage. *Critical Care*, 20(277), 2016. Publisher: Springer.
- [29] Friedman, Jerome H. and Hastie, Trevor and Tibshirani, Rob. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1):1–22, 2010.
- [30] Galles, David and Pearl, Judea. Testing Identifiability of Causal Effects. *arXiv*, February 2013.
- [31] Erdun Gao, Ignavier Ng, Mingming Gong, Wei Huang, Tongliang Liu, Kun Zhang, and Howard Bondell. MissDAG: Causal Discovery in the Presence of Missing Data with Continuous Additive Noise Models. *arXiv*, 2022. Publisher: arXiv.
- [32] Clark Glymour, Kun Zhang, and Peter Spirtes. Review of Causal Discovery Methods Based on Graphical Models. *Statistical Genetics and Methodology*, 10(524), 2019. Publisher: Frontiers in Genetics.
- [33] Greenland, Sander and Robins, James M. and Pearl, Judea. Confounding and Collapsibility in Causal Inference. *Statistical Science*, 14(1):29–46, 1999.
- [34] Tom M. and Tilling Kate Groenwold, Rolf H H. and Palmer. To Adjust or Not to Adjust? When a “Confounder” Is Only Measured After Exposure. *Epidemiology*, 32(2), March 2021.

- [35] Grossi, Giulio and Mattei, Alessandra and Papadogeorgou, Georgia. SMaC: Spatial Matrix Completion Method. *SIS 2023*, 2023.
- [36] Guido W. Imbens and Donald B. Rubin. *CAUSAL INFERENCE for Statistics, Social, and Biomedical Sciences: An Introduction*. 2015.
- [37] Alexander Hapfelmeier, Torsten Hothorn, and Kurt Ulm. Random Forest variable importance with missing data. *Universitätsbibliothek der Ludwig-Maximilians-Universität München*, 2012.
- [38] Alexander Hapfelmeier and Kurt Ulm. Variable selection by Random Forests using data with missing values. *Computational Statistics and Data Analysis*, 2014.
- [39] Hartwig, Fernando Pires and Tilling, Kate and Smith, Davey George. Empirically assessing the plausibility of unconfoundedness in observational studies. *arXiv*, February 2024.
- [40] Hastie, Trevor and Taylor, Jonathan and Tibshirani, Robert and Walther, Guenther. Forward stagewise regression and the monotone lasso. *Electronic Journal of Statistics*, 1:1–29, 2007.
- [41] Hastie, Trevor and Tibshirani, Robert and Wainwright, Martin. *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman and Hall/CRC, 1st edition edition, May 2015.
- [42] Jason S. Haukoos and Roger J. Lewis. The Propensity Score. *Journal of the American Medical Association*, 314(15):1637–1638, 2015. Publisher: American Medical Association.
- [43] Hazelwinkel, Audinga-Dea. Identifying and adjusting for bias due to missing outcome values in randomised controlled trials. *Open Access*.
- [44] Hernán, Miguel A. and Robins, James M. Instruments for Causal Inference: An Epidemiologist’s Dream? *Epidemiology*, 17(4):360–372, July 2006.
- [45] Hernán, Miguel A. and Robins, James M. *Causal Inference: What If*. Chapman & Hall / CRC, December 2020.
- [46] Jennifer L. Hill. Bayesian Nonparametric Modeling for Causal Inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011. Publisher: American Statistical Association, Institute of Mathematical Statistics, and Interface Foundation of North America.
- [47] Liang Hong, Mengqi Luo, Ruixue Wang, Peixin Lu, Wei Lu, and Long Lu. Big Data in Health Care: Applications and Challenges. *Data and Information Management Open Access*, 2018. Publisher: Sciendo.

- [48] Hu, Liangyuan and Joyce Lin, Jung-Yi and Ji, Jiayi. Variable selection with missing data in both covariates and outcomes: Imputation and machine learning. *Statistical Methods in Medical Research*, 30(12):2651–2671, 2021.
- [49] Huitfeldt, Anders and Stensrud, Mats J. and Suzuki, Etsuji. On the collapsibility of measures of effect in the counterfactual causal framework. *Emerging Themes in Epidemiology*, 16(1), 2019.
- [50] Höfler, M. Causal inference based on counterfactuals. *BMC Medical Research Methodology*, 5(28), 2005.
- [51] Kyung In Kim and Mark A van de Wiel. Effects of dependence in high-dimensional multiple testing problems. *BMC Bioinformatics*, 9(114), 2008. Publisher: BioMed Central Ltd.
- [52] Ip et al. Long-term cardiovascular safety of COVID-19 vaccination according to brand, dose and combinations: Cohort study of 46 million adults in England. *medRxiv*, 2024.
- [53] Dominik Janzing, Jonas Peters, Joris Mooij, and Bernhard Schölkopf. Identifying confounders using additive noise models. *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI2009)*, 2012. Publisher: arXiv.
- [54] Jeong, Hyunchai and Tian, Jin and Bareinboim, Elias. Finding and Listing Front-door Adjustment Sets. *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, 2022.
- [55] Jia, Zeyang and Imai, Kosuke and Li, Michael Lingzhi. The Cram Method for Efficient Simultaneous Learning and Evaluation. *arXiv*, March 2024.
- [56] Joffe, Marshall M. and Yang, Wei Peter and Feldman, Harold I. Selective Ignorability Assumptions in Causal Inference. *The International Journal of Biostatistics*, 6(2):11, 2010.
- [57] Kacprzyk, Krzysztof and Holt, Samuel and Berrevoets, Jeroen and Qian, Zhaozhi and van der Schaar, Mihaela. ODE Discovery for Longitudinal Heterogeneous Treatment Effects Inference. *ICLR*, 2024.
- [58] Kalos, Malvin H. and Whitlock, Paula A. *Monte Carlo Methods. Second Edition*.
- [59] Kaltiala-Heino, R and Bergman, H and Työläjärvä, M and Frisé, L. Gender dysphoria in adolescence: current perspectives. *Adolescent Health, Medicine and Therapeutics*, 2018(9):31–41, 2018.
- [60] Kaltiala, R. and Heino, Elias and Työläjärvä, M and Suomalainen, L. Adolescent development and psychosocial functioning after starting cross-sex hormones for gender dysphoria. *Nordic Journal of Psychiatry*, 74(3):213–219, 2020.

- [61] Edward H. Kennedy. Towards Optimal Doubly Robust Estimation of Heterogeneous Causal Effects. *Department of Statistics & Data Science Carnegie Mellon University*, 2020. Publisher: Carnegie Mellon University.
- [62] Koehler, Elizabeth and Brown, Elizabeth and Haneuse, Sebastien J-P A. On the Assessment of Monte Carlo Error in Simulation-Based Statistical Analyses. *American Statistical Association*, 63(2):155–162, May 2009.
- [63] Kong, Dehan and Yang, Shu and Wang, Linbo. Identifiability of causal effects with multiple causes and a binary outcome. *Biometrika*, 109(1):265–272, March 2022.
- [64] Soren R. Kunzel, Jasjeet S. Sekhon, Peter J. Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116(10):4156–4165, 2019. Publisher: PNAS.
- [65] Kyung, Minjung and Gill, Jeff and Ghosh, Malay and Casella, George. Penalized Regression, Standard Errors, and Bayesian Lassos. *Bayesian Analysis*, 5(2):369–412, 2010.
- [66] Labrecque, Jeremy and Swanson, Sonja A. Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr Epidemiol Rep*, 5:214–220, 2018.
- [67] Lai, Benjamin and Yu, Huang-Pin and Change, Yu-Ling and Wang, Liang-Chin and Chen, Che-Kai and Zhang, Weiya and Doherty, Michael and Chang, Shang-Hung and Hsu, Jun-Te and Yu, Kuang-Hui and Kuo, Chang-Fu. Assessing the causal relationships between gout and hypertension: a bidirectional Mendelian randomisation study with coarsened exposures. *Arthritis Research & Therapy*, 24(243), 2022.
- [68] Michael T. Lawton and G. Edward Vates. Subarachnoid Hemorrhage. *N Engl J Med*, 377:257–266, 2017. Publisher: Massachusetts Medical Society.
- [69] Lee, Paul H. Should we adjust for a confounder if empirical and theoretical criteria yield contradictory results? A simulation study. *Scientific Reports*, 4(6085), 2014.
- [70] Lee, Paul H. and Burstyn, Igor. Identification of confounder in epidemiologic data contaminated by measurement error in covariates. *BMC Medical Research Methodology*, 16(54), 2016.
- [71] Jonathan Levy, Mark van der Laan, Alan Hubbard, and Romain Pirracchio. A fundamental measure of treatment effect heterogeneity. *Journal of Causal Inference*, 9:83–108, 2021. Publisher: De Gruyter.

- [72] Linderman, George C. and Rachh, Manas and Hoskins, Jeremy G. and Steinberger, Stefan and Kluger, Yuval. Efficient Algorithms for t-distributed Stochastic Neighborhood Embedding. *arXiv*, December 2017.
- [73] Lippe, Phillip and Magliacane, Sara and Löwe, Sindy and Asano, Yuki M and Cohen, Taco and Gavves, Efstratios. BISCUIT: Causal Representation Learning from Binary Interactions. *PMLR*, 216:1263–1273, 2023.
- [74] Little, Charles H. C. *Combinatorial Mathematics V: Proceedings of the Fifth Australian Conference, Held at the Royal Melbourne Institute of Technology*, volume 622 of *Lecture Notes in Mathematics (LNM)*. Springer Berlin, Heidelberg, August 1976.
- [75] Loh, Wen Wei and Vansteelandt, Stijn. Confounder selection strategies targeting stable treatment effect estimators. *Statistics in Medicine*, 40(3):607–630, 2020.
- [76] Lousdal, Mette Lise. An introduction to instrumental variable assumptions, validation and estimation. *Lousdal Emerg Themes Epidemiol*, 15(1), 2018.
- [77] Lynn, Peter and Kaminska, Olena. Weighting Strategy for Understanding Society. Technical report, November 2010.
- [78] MacKinnon, David P. and Fairchild, Amanda J. and Fritz, Matthew S. Mediation Analysis. *Annual Review of Psychology*, 58:593–614, 2007.
- [79] Maldonado, George and Greenland, Sander. Simulation Study of Confounder-Selection Strategies. *American Journal of Epidemiology*, 138(11):923–936, 1993.
- [80] Scott Malec. Using the Literature to Identify Confounders. *UT SBMI Dissertations (Open Access)*, 41, 2018. Publisher: University of Texas Health Science Center.
- [81] Scott A. Malec, Peng Wei, Elmer V. Bernstam, Richard D. Boyce, and Trevor Cohen. Using computable knowledge mined from the literature to elucidate confounders for EHR-based pharmacovigilance. *Journal of Biomedical Informatics*, 117, 2021. Publisher: Elsevier Inc.
- [82] Many. Impact of vaccination on the association of COVID-19 with arterial and venous thrombotic diseases: an OpenSAFELY cohort study using linked electronic health records. *Pre-print*.
- [83] Many. Implementing Germ Defence digital behaviour change intervention via all primary care practices in England to reduce respiratory infections during the COVID-19 pandemic: an ecient cluster randomised controlled trial using the OpenSAFELY platform. *Pre-print*.



- [84] Many. MIMIC-III, a freely accessible critical care database. *Scientific Data*, 3(160035), 2016.
- [85] Many. Control of Confounding and Reporting of Results in Causal Inference Studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. *Ann Am Thorac Soc*, 16(1):22–28, January 2019.
- [86] Many. *The 58th Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference*. July 2020.
- [87] Many. Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example\*. *Critical Care Medicine*, 49(6):e563–e577, June 2021.
- [88] Many. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *International Journal of Epidemiology*, pages 620–632, 2021.
- [89] Many. Facilitators and barriers to community pharmacy PrEP delivery: A scoping review. *medRxiv*, 2023.
- [90] Many. Simulation Studies for Methodological Research in Psychology: A Standardized Template for Planning, Preregistration, and Reporting. *PsyArXiv*, 2023.
- [91] Many. A TRIAL EMULATION COMPARING THROMBOPROPHYLAXIS VERSUS NO THROMBOPROPHYLAXIS AMONG WARD-BASED PATIENTS WITH COVID-19. *Heart*, 109:A1–A324, 2023.
- [92] Many. Prostate-Specific Antigen Screening and 15-Year Prostate Cancer Mortality: A Secondary Analysis of the CAP Randomized Clinical Trial. *Journal of the American Medical Association*, April 2024.
- [93] Many. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). *Environment International*, 186:108–602, 2024.
- [94] Many. Use of artificial intelligence in critical care: opportunities and obstacles. *Critical Care*, 28(113), 2024.
- [95] Maya B. Mathur. The M-Value: A Simple Sensitivity Analysis for Bias Due to Missing Data in Treatment Effect Estimates. *American Journal of Epidemiology*, 192(4), December 2022.
- [96] Maya B. Mathur and Ilya Shpitser. Simple graphical rules to assess selection bias in general-population and selected-sample treatment effects. *American Journal of Epidemiology*.

- [97] Maya B. Mathur and Ilya Shpitser and Tyler J. VanderWeele. A common-cause principle for eliminating selection bias in causal estimands through covariate adjustment. *OSF Preprints*, 2023.
- [98] Gary C. McDonald. Ridge Regression. *WIREs Computational Statistics*, 1(1):93–100, 2009. Publisher: Wiley.
- [99] R McNamee. Confounding and Confounders. *Occp Environ Med*, 60:227–234, 2003. Publisher: BMJ Journals.
- [100] Christopher J. McWilliams, Daniel J. Lawson, Raul Santos-Rodriguez, Iain D. Gilchrist, Alan Champneys, Timothy H. Gould, Mathew JC Thomas, and Christopher P. Bourdeaux. Towards a decision support tool for intensive care discharge: machine learning algorithm development using electronic healthcare data from MIMIC-III and Bristol, UK. *BMJ Open*, 9(3), 2019. Publisher: British Medical Journal Publishing Group. eprint: <https://bmjopen.bmj.com/content/9/3/e025925.full.pdf>.
- [101] Milliard, Louise A. C. and Smith, George Davey and Tilling, Kate. Using the global randomization test as a Mendelian randomization falsification test for the exclusion restriction assumption. *European Journal of Epidemiology*, 2024, 2024.
- [102] Miot, HA. Analysis of ordinal data in clinical and experimental studies. *J Vasc Bras*, 19, 2020.
- [103] Pablo Morales-Alvarez, Angus Lamb, Simon Woodhead, Simon Peyton Jones, Miltos Allamanis, and Cheng Zhang. VICAUSE: Simultaneous missing value imputation and causal discovery. Technical Report MSR-TR-2021-14, ICML 2021 workshop on the Neglected Assumptions in Causal Inference, July 2021.
- [104] Morris, Tim P and White, Ian R and Crowther, Michael J. Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11):2074–2102, January 2019.
- [105] Michael C. Neale, Michael D. Hunter, Joshua N. Pritikin, Mahsa Zahery, Timothy R. Brick, Robert M. Kirkpatrick, Ryne Estabrook, Timothy C. Bates, Hermine H. Maes, and Steven M. Boker. OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 81(2):535–549, 2016. Publisher: Psychometric Society.
- [106] Nyberg, Joakim and Jonsson, E. Niclas and Karlsson, Mats O. and Häggström, Jonas. Properties of the full random-effect modeling approach with missing covariate data. *Statistics in Medicine*, pages 1–18, 2023.
- [107] Ogburn, Elizabeth L. and VanderWeele, Tyler J. Bias attenuation results for nondifferentially mismeasured ordinal and coarsened confounders. *Biometrika*, 100(1):241–248, 2013.

- [108] Osborne, Michael R. and Presnell, Brett and Turlach, Berwin A. On the LASSO and its Dual. *Journal of Computational and Graphical Statistics*, 9(2):319–337, 2000.
- [109] Park, Chan and Richardson, David and chetgen Tchetgen, Eric J. Single Proxy Control. *arXiv*, May 2024.
- [110] Park, Junhyung and Buchholz, Simon and Schölkopf, Bernhard and Muandet, Krikamol. A Measure-Theoretic Axiomatisation of Causality. *NeurIPS*, 2023.
- [111] Pearl, Judea. Causal diagrams for empirical research. *Biometrika*, 82(4):669–710, 1995.
- [112] Pearl, Judea. Causality: Models, Reasoning and Inference. *Econometric Theory*, 19:675–685, 2003.
- [113] Pearl, Judea. Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146, 2009.
- [114] Pearl, Judea. An Introduction to Causal Inference. *The International Journal of Biostatistics*, 6(2), 2010.
- [115] Pearl, Judea. The Seven Tools of Causal Inference, with Reflections on Machine Learning. *Communications of the ACM*, 62(3):54–60, 2019.
- [116] Pearl, Judea and Geffner, Hector and Dechter, Rina and Halpern, Joseph Y. *Probabilistic and Causal Inference: The Works of Judea Pearl*, volume 36. 1 edition, March 2022.
- [117] Pearl, Judea and Glymour, Madelyn and Jewell, Nicholas P. *Causal Inference in Statistics: A Primer*. Wiley, March 2016.
- [118] Jonas Peters, Joris M. Mooji, Dominik Janzing, and Bernhard Scholkopf. Causal Discovery with Continuous Additive Noise Models. *Journal of Machine Learning Research*, 15:2009–2053, 2014. Publisher: JMLR.
- [119] Klearchos Psychogios and Georgios Tsivgoulis. Subarachnoid Hemorrhage, Vasospasm, and Delayed Cerebral Ischemia. 2019.
- [120] Qian, Wei and Yang, Yuhong. Model selection via standard error adjusted adaptive lasso. *Ann Inst Stat Math*, 65:295–318, 2013.
- [121] Qian, Zhaozhi and Kacprzyk, Krzysztof and van der Schaar, Mihaela. D-CODE: Discovering Closed-form ODEs from Observed Trajectories. *ICLR*, 2022.
- [122] R Core Team. *R Documentation: Random Number Generation*. R Foundation for Statistical Computing, Vienna, Austria, 2021.

- [123] R White, Ian and Pham, Tra My and Quartagno, Matteo and Morris, Tim P. How to check a simulation study. *International Journal of Epidemiology*, 00:1–7, 2023.
- [124] Jeremy A. Rassen, M. Alan Brookhart, Robert J. Glynn, Murray A. Mittelman, and Sebastian Schneeweiss. Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. *Journal of Clinical Epidemiology*, 62:1226–1232, 2009. Publisher: Elsevier.
- [125] Rehkopf, David H. and Glymour, Maria M. and Osypuk, Theresa L. The Consistency Assumption for Causal Inference in Social Epidemiology: When a Rose is Not a Rose. *Curr Epidemiol Rep*, 3(1):63–71, March 2016.
- [126] Reid, Stephen and Tibshirani, Robert and Friedman, Jerome. A STUDY OF ERROR VARIANCE ESTIMATION IN LASSO REGRESSION. *Statistica Sinica*, 26(1):35–67, 2016.
- [127] Alexander G. Reisach, Christof Seiler, and Sebastian Weichwald. Beware of the Simulated DAG! Causal Discovery Benchmarks May Be Easy To Game, 2021.
- [128] Renaud, Olivier and Victoria-Feser, Maria-Pia. A robust coefficient of determination for regression. *Journal of Statistical Planning and Inference*, 140:1852–1862, 2010.
- [129] Robinson, R. W. Counting unlabeled acyclic digraphs. *Combinatorial Mathematics V. Lecture Notes in Mathematics*, 622, 1977.
- [130] Rosenbaum, Paul R. and Rubin, Donald B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.
- [131] Rothman, Kenneth J. et al. Serum Beta-Carotene: A Mechanism or "Yellow Finger"? *Epidemiology*, 3(4):277–279, July 1992.
- [132] Ruuska et al. All- cause and suicide mortalities among adolescents and young adults who contacted specialised gender identity services in Finland in 1996–2019: a register study. *British Medical Journal*, 2024(27):1–6, 2024.
- [133] Saarela, Olli and Stephens, David A. and Moodie, Erica E. M. The Role of Exchangeability in Causal Inference. *Statistical Science*, 2023.
- [134] Luca Salmasi and Enrice Capobianco. Use of instrumental variables in electronic health record-driven models. *Statistical Methods in Medical Research*, 27(2):608–621, 2018. Publisher: SAGE Journals.

- [135] Schooling, C. Mary and Jones, Heidi E. Clarifying questions about “risk factors”: predictors versus explanation. *Emerging Themes in Epidemiology*, 15(10), 2018.
- [136] Schuler, Megan S. and Rose, Sherri. Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. *American Journal of Epidemiology*, 185(1):65–73, January 2017.
- [137] Seedat, Nabeel and Imrie, Fergus and Bellot, Alexis and Qian, Zhaozhi and van der Schaar, Mihaela. Continuous-Time Modeling of Counterfactual Outcomes Using Neural Controlled Differential Equations. *ICML*, 2022.
- [138] Shen, Xinwei and Bühlmann, Peter and Taeb, Armeen. Causality-oriented robustness: exploiting general additive interventions. *arXiv*, 2023.
- [139] Susan M. Shortreed and Ashkan Ertefaie. Outcome-Adaptive Lasso: Variable Selection for Causal Inference. *Biometrics*, 73:1111–1122, 2017. Publisher: Wiley.
- [140] Natalia da Silva, Dianne Cook, and Eun-Kyung Lee. Interactive graphics for visually diagnosing forest classifiers in R. *Computational Statistics*, 2023. Publisher: Springer Nature.
- [141] Sloane, Neil J. A. and The OEIS Foundation Inc. A003024: Number of acyclic digraphs (or DAGs) with n labeled nodes, 2024.
- [142] Smith, George Davey and Phillips, Andrew N. Confounding in epidemiological studies: why “independent” effects may not be all they seem. *British Medical Journal*, 305(6856):757–759, September 1992.
- [143] Swanson, Sonja A. and Miller, Matthew and Robins, James M. and Hernán, Miguel A. Definition and Evaluation of the Monotonicity Condition for Preference-Based Instruments. *Epidemiology*, 26(3):414–420, May 2015.
- [144] Victoria Kubuta Talbot, Denis and Massamba. A descriptive review of variable selection methods in four epidemiologic journals: there is still room for improvemen. *European Journal of Epidemiology*, 34:725–730, 2019.
- [145] Talbot, Denis and Diop, Awa and Lavigne-Robichaud, Mathilde and Brisson, Chantal. The change in estimate method for selecting confounders: A simulation study. *Statistical Methods in Medical Research*, 30(9):2032–2044, 2021.
- [146] Emma Tarmey. ICCA Vasospasm Predictive Model, September 2023.
- [147] Emma Tarmey. Variable Selection Simulation Study, May 2023.

- [148] Peter W G Tennant, Eleanor J Murray, Kellyn F Arnold, Laurie Berrie, Matthew P Fox, Sarah C Gadd, Wendy J Harrison, Claire Keeble, Lysie R Ranker, Johannes Textor, Georgia D Tomova, Mark S Gilthorpe, and George T H Ellison. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *International Journal of Epidemiology*, 50(2):620–632, December 2020. .eprint: <https://academic.oup.com/ije/article-pdf/50/2/620/37947634/dyaa213.pdf>.
- [149] Laurien L. Teunissen, Gabriel J. E. Rinkel, Ale Algra, and J. van Gijn. Risk Factors for Subarachnoid Hemorrhage: A Systematic Review. *Stroke*, 27:544–549, 1996. Publisher: American Heart Association.
- [150] Thomas S. Richardson and James M. Robins. Single World Intervention Graphs (SWIGs): A Unification of the Counterfactual and Graphical Approaches to Causality. *Working Paper*, 128, April 2013.
- [151] Tibshirani, Robert. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, 58(1):267–288, 1996.
- [152] Tipton, Elizabeth and Mamakos, Michalis. Designing Randomized Experiments to Predict Unit-Specific Treatment Effects. *arXiv*, 2023.
- [153] Ruibo Tu, Kun Zhang, Paul Ackermann, Bo Christer Bertilson, Clark Glymour, Hedvig Kjellstrom, and Cheng Zhang. Causal Discovery in the Presence of Missing Data. *arXiv*, 2018. Publisher: arXiv.
- [154] Tudball, Matthew J. and Bowden, Jack and Hughes, Rachael A. and Ly, Amanda and Munafò, Marcus R. and Tilling, Kate and Zhao, Qingyuan and Davey Smith, George. Mendelian randomisation with coarsened exposures. *Genetic Epidemiology*, 45:338–350, 2021.
- [155] Tyler J. VanderWeele and Ilya Shpitser. A new criterion for confounder selection. *Biometrics*, 67(4):1406–1413, December 2011.
- [156] VanderWeele, Tyler J. Mediation Analysis: A Practitioner’s Guide. *Annual Review of Public Health*, 37:17–32, 2016.
- [157] VanderWeele, Tyler J. Principles of Confounder Selection. *European Journal of Epidemiology*, 34:211–219, 2019.
- [158] Greg Vaughan. Stagewise Estimating Equations. *University of Connecticut*, 2017. Publisher: University of Connecticut.
- [159] TS Verma and Judea Pearl. Equivalence and Synthesis of Causal Models. In *Probabilistic and Causal Inference: The Works of Judea Pearl*, pages 221–236. Association for Computing Machinery, New York, NY, USA, 1 edition, 2022.

- [160] Stefan Wager and Susan Athey. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018. Publisher: American Statistical Association.
- [161] Henning Walter, Stefan and Tiemeier. Variable selection: current practice in epidemiological studies. *European Journal of Epidemiology*, 24:733–736, 2009.
- [162] Pengfei Wei, Zhenzhou Lu, and Jingwen Song. Variable importance analysis: A comprehensive review. *Reliability Engineering and System Safety*, 142:399–432, 2015. Publisher: Elsevier.
- [163] Westfall, Jacob and Yarkoni, Tai. Statistically Controlling for Confounding Constructs Is Harder than You Think. *PLoS ONE*, 11(3), 2016.
- [164] White, Ian R. simsum: Analyses of simulation studies including Monte Carlo error. *The Strata Journal*, 10(3):369–385.
- [165] White, Ian R and Pham, Tra My and Quartagno, Matteo and Morris, Tim P. How to check a simulation study. *International Journal of Epidemiology*, 00:1–7, 2023.
- [166] Brian D. Williamson, Peter B. Gilbert, Marco Carone, and Noah Simon. Nonparametric variable importance assessment using machine learning techniques. *Biometrics*, 77:9–22, 2021. Publisher: The International Biometric Society.
- [167] Janine Witte, Ronja Foraita, and Vanessa Didelez. Multiple imputation and test-wise deletion for causal discovery with incomplete cohort data. *arXiv*, 2021. Publisher: arXiv.
- [168] Wu, Yinxiang and Kang, Hyunseung and Ye, Ting. Debiased Multivariable Mendelian Randomization. *arXiv*, February 2024.
- [169] Huiliang Xie and Jian Huang. SCAD-Penalised Regression in High-Dimensional Partially Linear Models. *The Annals of Statistics*, 37(2):673–696, 2009. Publisher: Institute of Mathematical Statistics.
- [170] Xie, Yan and Xu, Evan and Bowe, Benjamin and Al-Aly, Ziyad. Long-term cardiovascular outcomes of COVID-19. *Nature Medicine*, 28:583–590, March 2022.
- [171] Yiu, Andrew and Fong, Edwin and Holmes, Chris and Rousseau, Judith. Semiparametric posterior corrections. *arXiv*.
- [172] Zawadski, Roy S. and Grill, Joshua D. and Gillen, Daniel L. and "and for the Alzheimer’s Disease Neuroimaging Initiative". Framework for estimating causal effects in observational settings: comparing confounder adjustment and instrumental variables. *BMC Medical Research Methodology*, 23(122), 2023.

- [173] Jiaming Zeng, Gensheimer, Daniel L. Rubin, Susan Anthey, and Ross D. Shachter. Uncovering interpretable potential confounders in electronic medical records. *Nature Communications*, 13(1014), 2022. Publisher: Nature.
- [174] Cun-Hui Zhang. Nearly Unbiased Variable Selection Under Minimax Concave Penalty. *The Annals of Statistics*, 38(2):894–942, 2010. Publisher: Institute of Mathematical Statistics.
- [175] Zhang, Jiaqi and Jennings, Joel and Zhang, Cheng and Ma, Chao. Towards Causal Foundation Model: on Duality between Causal Inference and Attention. *arXiv*, September 2023.
- [176] Zhu, Yaqian and Hubbard, Rebecca A. and Chubak, Jessica and Roy, Jason and Mitra, Nandita. Core Concepts in Pharmacoepidemiology: Violations of the Positivity Assumption in the Causal Analysis of Observational Data: Consequences and Statistical Approaches. *Pharmacoepidemiol Drug Saf.*, 30(11):1471–1485, 2021.
- [177] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, 67(2):301–320, 2005. Publisher: Royal Statistical Society.
- [178] Zou, Hui. The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association*, 101(476):1418–1429, 2006.
- [179] Zuo, Shuozhi and Ghosh, Debashis and Ding, Peng and Yang, Fan. Mediation analysis with the mediator and outcome missing not at random. *arXiv*, September 2023.