

# EPI 953 Analytical Strategies for Observational Studies

Zhehui Luo

Fall 2021

## Contents

<b>1</b>	<b>Course Information</b>	<b>2</b>
1.1	Course Description	2
1.2	Course Objectives	2
1.3	Prerequisites	2
1.4	Textbook & Course Materials	2
<b>2</b>	<b>Course Outline/Schedule</b>	<b>3</b>
2.1	Planned Schedule	3
2.2	Other Notable Dates	5
2.3	Technical Assistance and Resources	5
<b>3</b>	<b>Grading Policy</b>	<b>5</b>
3.1	Graded Course Activities	5
3.2	Late Work Policy	6
3.3	Viewing Grades	6
3.4	Letter Grade Assignment (Grading Scale)	6
<b>4</b>	<b>Other Course Policies</b>	<b>6</b>
4.1	University Policies	6
4.2	Commit to Integrity: Academic Honesty	7
4.3	Inform Me of Any Accommodations Needed	7
4.4	Limits to Confidentiality	7
4.5	Commercialized Lecture Notes	7
4.6	Disruptive Behavior	7
4.7	Attendance	8
4.8	Build Rapport	8

# 1 Course Information

Logistics:

- Meeting Day/Time: MW 10:20-11:40
- Meeting location: E111 West Fee Hall
- Website: <https://d2l.msu.edu/d2l/home/1382023>

## Instructor Information

Instructor: Zhehui Luo

Personal pronouns: she, her, hers

Office: B627 West Fee Hall

Office Hours: Friday 11:00-12:00 via <https://msu.zoom.us/my/meet.zluo>

Appointments Outside Office Hours: By email.

Office Telephone: 517-884-3966. Email is the preferred mode of communication. E-mail: [zluo@msu.edu](mailto:zluo@msu.edu).

## 1.1 Course Description

The course is designed for students who have some experience with applied statistics in the social sciences or health sciences, especially epidemiology. It will cover methods for addressing measured and unmeasured confounding in observational (non-experimental or quasi-experimental) studies. Topics will include: counterfactual framework and causal diagram, comparisons of methods for measured confounding (regression, stratification, standardization and propensity score matching methods; marginal structural model; time-varying treatment and covariates; mediation analysis); unmeasured confounding (instrumental variables in linear and nonlinear models, bias and sensitivity analysis); and special program evaluation designs (difference-in-differences, synthetic control, regression discontinuity). The course does not cover causal inference for survival analysis; and is not a course for statistics theories.

## 1.2 Course Objectives

The ultimate goal of the course is to help you do better empirical research. From the definition of causal estimands to the interpretation of results, every step in research needs critical thinking. This course provides some tools to assist the process. Because the course is designed for PhD students in epidemiology (typically in the 3rd year), I expect you to have a research topic of interest and the final project for the course requires you to apply at least two things you learned in this course and reflects on whether and how the course helps your research.

## 1.3 Prerequisites

EPI 826 or EPI 826B are the official prerequisites for the course. They are graduate-level introductory courses in applied statistics, including linear and generalized linear regressions. Basic understanding of probability, independence, expectation and variance of functions of random variables is required. The statistical software for illustrations may be Stata, SAS or R. Program codes are indicated with the **Sans Serif** font. You are free to use any software in which you are proficient.

## 1.4 Textbook & Course Materials

There is no required textbook, but the free online book by [Hernán & Robins \(2020\)](#) will be the primary resource. Readings will come largely from journal articles in the biostatistics and epidemiology literature. More in-depth readings are in many excellent books ([Pearl, 2009b](#); [Imbens & Rubin, 2015](#); [Morgan & Winship, 2007](#); [VanderWeele, 2015](#); [Rosenbaum, 2002](#); [Lash et al., 2009](#); [Guo & Fraser, 2015](#); [Hong, 2015](#)), and a lay-person readable introduction by [Pearl and MacKenzie \(2018\)](#) is actually fun. You are expected to have read the required readings before each lecture (see [2.1](#)). Lecture notes, slides, data and programs will be distributed on D2L.

## 2 Course Outline/Schedule

The primary learning objectives are

- To obtain a deeper understanding of the relationship between statistical modeling and causal inference
- To better understand cutting-edge research methodologies
- To use at least some of the tools to improve the quality of your own research

### 2.1 Planned Schedule

**9/1** Is counterfactual language a new language? We will learn the formalism of the counterfactual framework for causal inference. The required reading is Chapter 1 of [Hernán & Robins \(2020\)](#). Supplemental readings include [Holland \(1986\)](#), [Pearl \(2009a, Section 2\)](#), [Imbens & Rubin \(2015, Chapters 1 and 2\)](#), and [Blalock Jr. \(1961, Chapter 1\)](#).

**9/6** Holiday - University Closed

**9/8** What is a directed acyclic graph (DAG) and what is a causal DAG. The required readings are [Greenland et al. \(1999a\)](#) and [Hernán & Robins \(2020, Chapter 6\)](#). Supplemental readings include [Shrier & Platt \(2008\)](#); [Textor et al. \(2016\)](#); [Elwert \(2013, Ch13\)](#); [Rothman et al. \(2008, Ch12\)](#).

**9/13** Yet another look at the concept of confounding. What is the back-door criterion? What is the relationship between regression adjustment and standardization (stratification)? The required readings are [Hernán & Robins \(2020, Chapter 7\)](#), [Greenland et al. \(1999b\)](#) and its follow up [Greenland & Robins \(2009\)](#). Supplemental readings include [Paik \(1985\)](#); [Shpitser et al. \(2010\)](#); [Shpitser & Tchetgen \(2016\)](#).

**9/15** Lab

**9/17** Homework 1 is due

**9/20** Propensity score methods, what is the big deal? We will take a look at the foundation paper of the method by [Rosenbaum & Rubin \(1983b\)](#) and try to understand why [Hernán & Robins \(2020\)](#) only briefly mentioned this in Chapter 15 of their book whereas [Imbens & Rubin \(2015\)](#) spend 7 chapters on the topic. Review and assessment of various methods by Peter Austin are very helpful ([2007a](#); [2007b](#); [2008](#); [2009](#); [2011](#); [2014](#); [2014](#); [2015b](#); [2015a](#); [2016](#); [2016](#)). Practical guidelines come out every day; use your judgement ([Granger et al., 2020](#); [Samuel et al., 2020](#)). Extensions to general settings are also interesting ([McCaffrey et al., 2013](#); [Granger et al., 2019](#); [Ling et al., 2020](#)). The book by [Guo & Fraser \(2015\)](#) is an excellent resource.

**9/22** Lab

**9/29** We will try to see how marginal structural models are different from familiar statistical models, how inverse probability weighting and standardization are two methods to estimate parameters of interest in these models, and how they are related to the parametric g-formula. The best description of these topics is included in Chapter 11 and 12 of [Hernán & Robins \(2020\)](#). Supplemental readings include [Robins et al. \(2000\)](#); [Sato & Matsuyama \(2003\)](#); [Mortimer et al. \(2005\)](#); [Cole & Hernán \(2008\)](#).

**10/1** Homework 2 is due

**10/4** Lab

**10/6** If you are familiar with the statistical models for longitudinal data analysis, this lecture may change how you think about them. Time-varying treatments, time-varying confounding and the g-estimation were designed for prolonged exposures ([Robins, 1986](#), this is the original paper for your reference, not a required reading); we will briefly examine Chapters 19, 20 and 21 of [Hernán & Robins \(2020\)](#). The general idea is summed up in [Bodnar et al. \(2004\)](#). A rare SAS book example gives the codes for a modified data ([Faries et al., 2010, Chapter 9](#)).

**10/13** Lab

10/15 Homework 3 is due

10/18 Mediation analysis has a long history and generated a lot of debate recently. We will briefly examine the differences and connections between the traditional and current approaches. Here are but a few examples: [VanderWeele \(2013\)](#); [VanderWeele & Vansteelandt \(2013\)](#); [Lange et al. \(2012\)](#); [Steen et al. \(2017\)](#). VanderWeele won the 2018 ASA Causality in Statistics Education award for his book on mediation and effect modification ([VanderWeele, 2015](#)).

10/20 Lab (Middle of semester)

10/25 Break Day (University open, class cancelled)

10/27 Midterm exam

11/1 Special designs for policy evaluations ([Imbens & Wooldridge, 2009](#)), including difference-in-differences (DID) and synthetic controls (SC), are becoming more popular in epidemiology and statisticians have made a lot of contributions recently. [This site](#) contains a rather complete description and references for DID and SC. We will not be able to cover everything; and perhaps will focus on the caveats ([Dimick & Ryan, 2014](#); [Ryan et al., 2015](#); [Ryan, 2018](#)).

11/3 Lab

11/5 Homework 4 is due

11/8 This part of the course focuses on unmeasured confounding and a class of methods to address that: instrumental variables (IV) methods. The key papers are [Angrist et al. \(1996\)](#); [Greenland \(2000\)](#); [Hernán & Robins \(2006\)](#). Because the IV methods have a long history in economics, it is perhaps not surprising all graduate level econometrics courses include the topic. Its popularity grew in epidemiology and medical research since the major study by McClellan et al. ([McClellan & Newhouse, 2000](#)), a special issue was published in Health Services Research ([McClellan & Newhouse, 2000](#)), and you may have heard of it under a different name, Mendelian randomization ([Didelez & Sheehan, 2007](#); [Burgess et al., 2011](#)). However, the method should not be used lightly ([Staiger & Stock, 1997](#); [Stock et al., 2002](#); [Didelez et al., 2010](#); [Wooldridge, 2016](#)).

11/10 Lab

11/15 When your research potentially contains unmeasured confounding, performing bias and bounds sensitivity (BBS) is always a good idea, to you and your readers. The book by [Lash et al. \(2009\)](#) is highly recommended. Other key papers include [Rosenbaum & Rubin \(1983a\)](#); [DiPrete & Gangl \(2004\)](#); [Nannicini \(2007\)](#); [Ding & VanderWeele \(2016\)](#).

11/17 Lab

11/19 Homework 5 is due

11/22 Generalizability and transportability, data fusion. These are very new topics and we will read together if we have time: [Bareinboim & Pearl \(2016\)](#); [Pearl & Bareinboim \(2019\)](#).

11/24 Lab

11/29 Putting everything together? Combination of experimental and non-experimental studies, emulate trials. What is the advantage of combining two camps together? These are uncharted areas and we will read together if we have time: [Hernán & Robins \(2020, Chapter 22\)](#)

12/1 Lab

12/3 Homework 6 is due 12/3

12/6 Student presentation

12/8 Student presentation

12/13-17 Finals week

## 2.2 Other Notable Dates

The last day to add this course is the end of the first week of classes. The last day to drop this course with a 100 percent refund and no grade reported is 9/27/2020. The last day to drop this course with no refund and no grade reported is 10/20/2020. You should immediately make a copy of your amended schedule to verify you have added or dropped this course.

9/3 Application for fall graduation deadline

9/8 Open add period ends

9/27 Tuition refund period ends

11/25-26 Holiday - University closed

12/12 Classes end

12/21 Grades due by 4:00 p.m. Grades available to students the following business day.

## 2.3 Technical Assistance and Resources

MSU has some very helpful assistance if you need (e.g., MSU Libraries [Discovery Services](#), Desire2Learn [Help](#)), and many free resources for our students (e.g., Electronic [databases and journals](#), EndNote Online [premium version](#)).

### D2L

For first time D2L users, log in with your MSU NetID and password at [d2l.msu.edu](http://d2l.msu.edu), select “Self-Registration” from the menu bar, register for and complete the course named “Students - Getting Started with D2L.” Also be aware of the “Help” option on the D2L Home page menu.

### Resource Center for Persons with Disabilities (RCPD)

- To make an appointment with a specialist, contact: (517) 353-9642 Or TTY: (517) 355-1293, or
- Register at [MYProfile.rcpd.msu.edu](http://MYProfile.rcpd.msu.edu)

## 3 Grading Policy

### 3.1 Graded Course Activities

The table below describes the graded course activities including points and activity description. The first column includes the points possible, and the second column includes a description for each activity. An important component of learning is active participation. Thus, the evaluation of the course includes Participation. We will have a discussion forum on D2L where you ask and answer each other’s questions and I will weigh in as necessary. We can co-write the Rubric for Homework, exam and final project during the semester.

Category	Points	Description
Participation	5	This includes in-class participation and online discussion forum presence.
Homework	60	There will be a bi-weekly homework (10 points each) with exercises using real datasets.
Midterm	15	In-class exam, open-book
Final Project	20	A presentation and report based on your research topic

### 3.2 Late Work Policy

Assignments for this course will be submitted electronically through D2L unless otherwise instructed. Assignments must be submitted by the given deadline or special permission must be requested from me before the due date. Extensions will not be given beyond the next assignment except under extreme circumstances. For late submission of homework and final project with permission, the following reduction of percentage points of a student's actual grade for that element will be taken: 1-day late 5%; 2-day late 10%, 3-day late 25%, more than 3 days late 50%. For late submission of homework and final project without permission, zero point will be given for that element.

### 3.3 Viewing Grades

Grades for homework and final exam will be available on D2L in 2 days. Grades for Participation and final project will be available on D2L on December 22.

### 3.4 Letter Grade Assignment (Grading Scale)

The table below describes the relationships between letter grades, percent, and performance. The first column describes the letter grade. The second column describes the percentage associated with that letter grade. The third column describes the performance represented by that letter grade and percentage.

Letter Grade	Percentage	Performance
4.0	90 to 100%	Excellent Work
3.5	80 to 89%	Very Good Work
3.0	70 to 79%	Above Average Work
2.5	60 to 69%	Mostly Average Work
2.0	50 to 59%	Below Average Work
1.5	40 to 49%	Poor Work
F	0 to 39%	Failing Work

## 4 Other Course Policies

Students are expected to adhere to the policies of MSU whether noted in this syllabus or not. While the below may appear at first glance to be common policy boilerplate there may be nuances or course specifics within it that the student must be aware of and adhere to.

### 4.1 University Policies

- All Michigan State University policies regarding academic integrity apply. I also endorse the Spartan Code of Honor and would love for you to make a pledge. [Spartan Code of Honor](#): “As a Spartan, I will strive to uphold values of the highest ethical standard. I will practice honesty in my work, foster honesty in my peers, and take pride in knowing that honor in ownership is worth more than grades. I will carry these values beyond my time as a student at Michigan State University, continuing the endeavor to build personal integrity in all that I do.”
- Mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. Services are available to assist you with addressing these and other concerns you may be experiencing. You can learn more about the broad range of confidential mental health services available on campus via the Counseling & Psychiatric Services (CAPS) website at [www.caps.msu.edu](http://www.caps.msu.edu).
- Religious Observance Policy: <https://reg.msu.edu/ROInfo/Notices/ReligiousPolicy.aspx>

## 4.2 Commit to Integrity: Academic Honesty

Article 2.III.B.2 of the [Academic Rights and Responsibilities](#) states that "The student shares with the faculty the responsibility for maintaining the integrity of scholarship, grades, and professional standards." In addition, the Department of Epidemiology and Biostatistics adheres to the policies on academic honesty as specified in [General Student Regulations 1.0, Protection of Scholarship and Grades](#). See [Spartan Life Online](#) for more information.

Therefore, unless authorized by me, you are expected to complete all course assignments, including homework, exams and final projects, without assistance from other people. You are expected to develop original work for this course; therefore, you may not submit course work you completed for another course to satisfy the requirements for this course. Students who violate MSU academic integrity rules may receive a penalty grade, including a failing grade on the assignment or in the course.

## 4.3 Inform Me of Any Accommodations Needed

According to the [Resource Center for Persons with Disabilities \(RCPD\)](#): MSU is committed to providing equal opportunity for participation in all programs, services and activities. Requests for accommodations by persons with disabilities may be made by contacting the RCPD at 517-884-RCPD or on the web at [rcpd.msu.edu](http://rcpd.msu.edu). Once your eligibility for an accommodation has been determined, you will be issued a Verified Individual Services Accommodation ("VISA") form. Please present this form to me at the start of the term and/or two weeks prior to the accommodation date (exam, project, etc.). Requests received after this date will be honored whenever possible.

## 4.4 Limits to Confidentiality

Materials submitted for this class are generally considered confidential pursuant to the University's student record policies. However, students should be aware that University employees, including instructors, may not be able to maintain confidentiality when it conflicts with their responsibility to report certain issues to protect the health and safety of MSU community members and others. As the instructor, I must report the following information to other University offices (including the Department of Police and Public Safety) if you share it with me:

- Suspected child abuse/neglect, even if this maltreatment happened when you were a child,
- Allegations of sexual assault or sexual harassment when they involve MSU students, faculty, or staff, and
- Credible threats of harm to you or to others.

These reports may trigger contact from a campus official who will want to talk with you about the incident that you have shared. In almost all cases, it will be your decision whether you wish to speak with that individual. If you would like to talk about these events in a more confidential setting you are encouraged to make an appointment with the [MSU Counseling Center](#).

## 4.5 Commercialized Lecture Notes

Commercialization of lecture notes and university-provided course materials is not permitted in this course. The Code of Teaching Responsibility requires instructors who permit students to commercialize their class lecture notes to include a statement in their course syllabi that gives such permission. Absent such permission, students may not do so.

## 4.6 Disruptive Behavior

Article 2.III.B.4 of [Student Rights and Responsibilities](#) for students at MSU states: "The student's behavior in the classroom shall be conducive to the teaching and learning process for all concerned." Article 2.III.B.10 states that "The student and the faculty share the responsibility for maintaining professional relationships based on mutual trust and civility." [General Student Regulation 5.02](#) states: "No student shall . . . obstruct, disrupt, or interfere with

the functions, services, or directives of the University, its offices, or its employees (e.g., classes, social, cultural, and athletic events, computing services, registration, housing and food services, governance meetings, and hearings).” Students whose conduct adversely affects the learning environment may be subject to disciplinary action through the Student Judicial Affairs office.

#### **4.7 Attendance**

University Attendance Policy (in part): Students whose names do not appear on the official class list for this course may not attend this class. Students who fail to attend the first four class sessions or the fifth class-day of the instruction, whichever occurs first, may be dropped from the course.

#### **4.8 Build Rapport**

If you find that you have any trouble keeping up with assignments or other aspects of the course, make sure to let me know as early as possible. As you will find, building rapport and effective relationships are key to becoming an effective professional. Make sure that you are proactive in informing me when difficulties arise during the semester so that I can help you find a solution.

## References

- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434), 444–455.
- Austin, P. C. (2008). The performance of different propensity-score methods for estimating relative risks. *Journal of Clinical Epidemiology*, 61(6), 537–45.
- Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*, 28(25), 3083–107.
- Austin, P. C. (2011). Comparing paired vs non-paired statistical methods of analyses when making inferences about absolute risk reduction in propensity-score matched samples. *Statistics in Medicine*, 30, 1292–1301.
- Austin, P. C. (2014). A comparison of 12 algorithms for matching on the propensity score. *Stat Med*, 33(6), 1057–69.
- Austin, P. C., Grootendorst, P., & Anderson (2007a). A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: a monte carlo study. *Statistics in Medicine*, 26(4), 734–53.
- Austin, P. C., Grootendorst, P., Normand, & Anderson (2007b). Conditioning on the propensity score can result in biased estimation of common measures of treatment effect: a monte carlo study. *Statistics in Medicine*, 26(4), 754–68.
- Austin, P. C., Jembere, N., & Chiu, M. (2016). Propensity score matching and complex surveys. *Stat Methods Med Res*.
- Austin, P. C. & Schuster, T. (2016). The performance of different propensity score methods for estimating absolute effects of treatments on survival outcomes: A simulation study. *Stat Methods Med Res*, 25(5), 2214–2237.
- Austin, P. C. & Small, D. S. (2014). The use of bootstrapping when using propensity-score matching without replacement: a simulation study. *Stat Med*, 33(24), 4306–19.
- Austin, P. C. & Stuart, E. A. (2015a). Estimating the effect of treatment on binary outcomes using full matching on the propensity score. *Stat Methods Med Res*.
- Austin, P. C. & Stuart, E. A. (2015b). Optimal full matching for survival outcomes: a method that merits more widespread use. *Stat Med*, 34(30), 3949–67.
- Bareinboim, E. & Pearl, J. (2016). Causal inference and the data-fusion problem. *Proc Natl Acad Sci U S A*, 113(27), 7345–52.
- Blalock Jr., H. M. (1961). *Causal Inferences in Nonexperimental Research*. The University of North Carolina Press.
- Bodnar, Davidian, Siega-Riz, & Tsiatis (2004). Marginal structural models for analyzing causal effects of time-dependent treatments: An application in perinatal epidemiology. *American Journal of Epidemiology*, 159.
- Burgess, S., Thompson, S. G., & CRP CHD Genetics Collaboration (2011). Avoiding bias from weak instruments in mendelian randomization studies. *International Journal of Epidemiology*, 40, 755–764.
- Cole, S. R. & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*, 168(6), 656–64.
- Didelez, V., Meng, S., & Sheehan, N. A. (2010). Assumptions of iv methods for observational epidemiology. *Statistical Science*, 25(1), 22–40.
- Didelez, V. & Sheehan, N. A. (2007). Mendelian randomization as an instrumental variable approach to causal inference. *Statistical Methods in Medical Research*, 16, 309–330.
- Dimick, J. B. & Ryan, A. M. (2014). Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA*, 312(22), 2401–2.
- Ding, P. & VanderWeele, T. J. (2016). Sensitivity analysis without assumptions. *Epidemiology*, 27(3), 368–77.

- DiPrete, T. & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology*, 34, 271–310.
- Elwert, F. (2013). *Handbook of Causal Analysis for Social Research*, (pp. 245–273). Springer.
- Faries, D. E., Obenchain, R. L., Haro, J. M., & Leon, A. C. (2010). *Analysis of Observational Health Care Data Using SAS*. SAS Institute Inc.
- Granger, E., Sergeant, J. C., & Lunt, M. (2019). Avoiding pitfalls when combining multiple imputation and propensity scores. *Stat Med*, 38(26), 5120–5132.
- Granger, E., Watkins, T., Sergeant, J. C., & Lunt, M. (2020). A review of the use of propensity score diagnostics in papers published in high-ranking medical journals. *BMC Med Res Methodol*, 20(1), 132.
- Greenland, S. (2000). An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology*, 29(4), 722–729.
- Greenland, S., Pearl, J., & Robins, J. M. (1999a). Causal diagrams for epidemiologic research. *Epidemiology*, 10(1), 37–48.
- Greenland, S. & Robins, J. M. (2009). Identifiability, exchangeability and confounding revisited. *Epidemiol Perspect Innov*, 6, 4.
- Greenland, S., Robins, J. M., & Pearl, J. (1999b). Confounding and collapsibility in causal inference. *Statistical Science*, 14(1), 29–46.
- Guo, S. & Fraser, M. W. (2015). *Propensity Score Analysis: Statistical Methods and Applications*. Sage Publications.
- Hernán, M. A. & Robins (2006). Instruments for causal inference: an epidemiologist’s dream? *Epidemiology*, 17(4), 360–72.
- Hernán, M. A. & Robins, J. M. (2020). *Causal Inference: What If*. <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>: Online.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945–960.
- Hong, G. (2015). *Causality in a Social World: Moderation, Mediation and Spill-over*. Wiley & Sons.
- Imbens, G. W. & Rubin, D. B. (2015). *Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction*. Cambridge University Press.
- Imbens, G. W. & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86.
- Lange, T., Vansteelandt, S., & Bekaert, M. (2012). A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol*, 176(3), 190–5.
- Lash, T. L., Fox, M. P., & Fink, A. K. (2009). *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer Science+Business Media, Inc.
- Ling, A., Montez-Rath, M., Mathur, M., Kapphahn, K., & Desai, M. (2020). How to apply multiple imputation in propensity score matching with partially observed confounders: a simulation study and practical recommendations. *Journal of Modern Applied Statistical Methods*, 19(1), eP3439.
- McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med*, 32(19), 3388–414.
- McClellan & Newhouse (2000). Overview of the special supplement issue "instrumental variables analysis: Applications in health services research.". *Health Services Research*, 35(5 Part II), 1061–1069.

- Morgan, S. L. & Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press.
- Mortimer, K. M., Neugebauer, R., van der Laan, M., & Tager, I. B. (2005). An application of model-fitting procedures for marginal structural models. *American Journal of Epidemiology*, 162(4), 382–388.
- Nannicini, T. (2007). Simulation-based sensitivity analysis for matching estimators. *Stata Journal*, 7, 334–350.
- Paik, M. (1985). A graphic representation of a three-way contingency table: Simpson’s paradox and correlation. *The American Statistician*, 39(1), 53–54.
- Pearl, J. (2009a). Causal inference in statistics: an overview. *Statistical Surveys*, 3, 96–146.
- Pearl, J. (2009b). *Causality: Models Reasoning, and Inference*. New York, NY: Cambridge University Press, 2nd edition.
- Pearl, J. & Bareinboim, E. (2019). Note on ”generalizability of study results”. *Epidemiology*, 30(2), 186–188.
- Pearl, J. & MacKenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York.
- Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7, 1393–1512.
- Robins, J. M., Hernán, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11, 550–560.
- Rosenbaum, P. R. (2002). *Observational Studies*. New York, NY: Springer, 2nd edition.
- Rosenbaum, P. R. & Rubin, D. B. (1983a). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society Series B-Methodological*, 45(2), 212–218.
- Rosenbaum, P. R. & Rubin, D. B. (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern Epidemiology*. Lippincott Williams and Wilkins.
- Ryan, A. M. (2018). Well-balanced or too matchy-matchy? the controversy over matching in difference-in-differences. *Health Serv Res*, 53(6), 4106–4110.
- Ryan, A. M., Burgess, Jr, J. F., & Dimick, J. B. (2015). Why we should not be indifferent to specification choices for difference-in-differences. *Health Serv Res*, 50(4), 1211–35.
- Samuel, M., Batomen, B., Rouette, J., Kim, J., Platt, R. W., Brophy, J. M., & Kaufman, J. S. (2020). Evaluation of propensity score used in cardiovascular research: a cross-sectional survey and guidance document. *BMJ Open*, 10(8), e036961.
- Sato, T. & Matsuyama, Y. (2003). Marginal structural models as a tool for standardization. *Epidemiology*, 14, 680–686.
- Shpitser, I. & Tchetgen, E. T. (2016). Causal inference with a graphical hierarchy of interventions. *Ann Stat*, 44(6), 2433–2466.
- Shpitser, I., VanderWeele, T. J., & Robins, J. M. (2010). On the validity of covariate adjustment for estimating causal effects. In P. Grünwald & P. Spirtes (Eds.), *UAI* (pp. 527–536): AUAI Press.
- Shrier, I. & Platt, R. W. (2008). Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology*, 8(70).
- Staiger, D. & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557–586.
- Steen, J., Loeys, T., Moerkerke, B., & Vansteelandt, S. (2017). Flexible mediation analysis with multiple mediators. *Am J Epidemiol*, 186(2), 184–193.

- Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, 20(4), 518–529.
- Textor, J., van der Zander, B., Gilthorpe, M. S., Liskiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: the r package 'dagitty'. *Int J Epidemiol*, 45(6), 1887–1894.
- VanderWeele, T. (2013). A three-way decomposition of a total effect into direct, indirect, and interactive effects. *Epidemiology*, 24(2), 224–232.
- VanderWeele, T. J. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press.
- VanderWeele, T. J. & Vansteelandt, S. (2013). Mediation Analysis with Multiple Mediators. *Epidemiol Method*, 2(1), 95–115.
- Wooldridge, J. M. (2016). Should instrumental variables be used as matching variables? *Research in Economics*, 70(2), 232–237.