

CURRICULUM VITAE and BIBLIOGRAPHY (August 2022)

CONTACT INFORMATION

Name: Mark Johannes van der Laan.

Nationality: Dutch.

Marital status: Married to Martine with children Laura, Lars, and Robin.

University address:

University of California
Division of Biostatistics
School of Public Health
108 Haviland Hall
Berkeley, CA 94720-7360
email: laan@stat.berkeley.edu.

Telephone number:

office: 510-643-9866
fax: 510-643-5163

Web address:

www.stat.berkeley.edu/laan
Working Papers, Division of Biostatistics: www.bepress.com/ucbbiostat

EDUCATION

1990-1993: Department of Mathematics, Utrecht University.

Ph.D student of Prof. Dr. R.D. Gill. Position included 25% teaching, 75% research and education.

Specialization in Estimation in Semiparametric and Censored Data Models.

1991-1992: **University of California, Berkeley.**

Statistics Program at M.S.R.I.: "Semiparametric Models and Survival Analysis".

Research with guidance by second Promotor Prof. Dr. P.J. Bickel.

Subject: "Efficient Estimation in the Bivariate Censoring Model".

December 13, 1993: Official Public Defense of Ph.D Thesis.

1985-1990: **Masters degree in Mathematics at the University of Utrecht, The Netherlands.** Statistics Major.

1988-1989: One year study, Masters degree courses at the Department of Statistics,
North Carolina State University, Raleigh,

North Carolina, U.S.A.

G.P.A 4.0, Dean's List.

1989-1990: Masters thesis under guidance of Prof. Dr. R.D. Gill.

Subject: The Dabrowska Estimator and the Functional Delta method.

Grade (from 1-10, 10=top): 9.5.

Official Completion: May 1, 1990.

ACADEMIC POSITIONS

- *2021-present: Member of Department of Computational Precision Medicine.
- *2018-present: Chair Group of Biostatistics, UC Berkeley.
- 2006-present: Jiann-Ping Hsu/Karl E. Peace Endowed Chair in Biostatistics.
- 2013-2018: Investigator and core leader of the methods workgroup of the Sustainable East African Research in Community Health (SEARCH).
- *2016-present: Academic Director of Center of Targeted Learning in Precision Health.
- September 2004-present: Director of the Biostatistics and Computing core of the Superfund Research Program (Genomics in environmental science) in the School of Public Health headed by Prof Martyn Smith.
- 2017- Director of core of biostatistics for environmental grant ECHO of Brenda Eskenazi.
- 2000-present: Professor Biostatistics and Statistics (joint appointment), School of Public Health and Department of Statistics, University of California, Berkeley.
- 1998-2000: Associate Professor of Biostatistics and Statistics (joint appointment), School of Public Health and Department of Statistics, University of California, Berkeley.
- 2005-2009: Long-term statistical consultant/adjunct professor for Bioinformatics at the Aging Buck Institute, Novato.
- Spring 2007: Miller Professor, UC Berkeley.
- 1994-1998: Assistant Professor, Biostatistics, School of Public Health, University of California, Berkeley.
- Summer and Fall 2004: Visiting Assistant Professor, Department of Mathematics, Free University, Amsterdam (Aad van der Vaart).
- 1994 Neyman Visiting Professor, Department of Statistics, UC Berkeley.
- 1990-1993: Teaching Assistant in the Department of Mathematics during Ph.D position, Utrecht, for masters degree courses.
 - Algebra B and C, Math. Analysis B and C (physics students).
 - Intr. Stochastic Analysis, Measure and Integration theory (math. students).
 - Mathematics I and II (chemistry students).
 - Statistics for Physical Sciences.

RESEARCH AREAS See <http://www.stat.berkeley.edu/~laan/Research/research.html>

My main research interests are

- Developing statistical methodology and theory for analyzing high dimensional **censored longitudinal data** from observational and randomized studies.
- Statistical methods for **causal inference** in longitudinal studies with both informative treatment assignment and informative censoring, clinical trials, safety analysis.
- Adaptive designs and surveillance systems.
- Statistical methods for the **analysis of genomic data** in computational biology and medical/epidemiological research.

Dependent network and time series data

The application of these methods in collaboration with scientists (**comparative effectiveness research**).

Highly Adaptive Lasso as new class of machine learning algorithms.

Adaptive designs for time-series.

Super learning and **Targeted Maximum Likelihood Estimation** (TMLE).

Targeted Learning: We have developed a new approach to statistical learning called targeted maximum likelihood learning. In general, based on this approach we develop machines that take out the human intervention in data analysis, and provide reliable robust and optimal estimators of the desired target estimand, with as honest assessment of uncertainty. The methods generalize machine learning to the fully automated learning of any kind of parameter of the data generating experiment. It incorporates the state of the art in adaptive estimation we termed super-learning.

Patent: B07-108 Targeted Maximum Likelihood Learning.

ACADEMIC AWARDS, GRANTS, HONORS

February 23, 1991: Scholarship of the VSB Foundation presented by the Chairman of the Board of Directors at the opening of the 355th anniversary of the University of Utrecht.

March 27, 1991:

Second Prize in the contest of the best (Netherlands) Masters thesis in Statistics or Operational Research (1989-1990) presented by the V.V.S.

Publication of a summary of the masters thesis in *Statistica Neerlandica*.

January, 1994 through June, 1994: Neyman Visiting Assistant Professor, Department of Statistics, University of California, Berkeley.

July 1994: **Martin Sisters Chair**, School of Public Health, UC Berkeley, 3 years (\$15,000 per year).

1995: Two faculty grants of \$3000 each, Committee on Research, UC Berkeley.

1995, 1996: Two junior faculty mentor grants, \$750 and \$1000, Office of the Chancellor, UC Berkeley.

1996-97: **Hellman Family Faculty Award**, \$20,000, Office of the Chancellor, UC Berkeley.

1996: Ph.D thesis selected to be published in book form (CWI-tract) by the Centre of Mathematics and Computer Science, Amsterdam.

June 1996: **FIRST Award, 5 year NIH grant**, 1996-2001, \$498,726 (total costs). Title: Locally Efficient Estimation with High Dimensional Data Structures. Score in top 7 %.

October-December, 1998: Visiting Professor, Department of Mathematics, Free University Amsterdam, host: Prof. Dr. A.W. van der Vaart.

June 1999: **NIAID Award, 3 year grant**, 1999-2002. \$357,000 (total costs).

October-December, 1999: Honorary Visiting Professor, Department of Statistics, Auckland, New Zealand, host: Prof. Dr. A. Scott. Score in top 13%. Title: Causal Inference and Longitudinal Aids Studies.

September 2000, 3 year grant: **LLNL 3-year grant**, 2000-2003. Title: “Statistical Inference from Microarray Data with Applications in Breast Cancer Research”.

September 2000, 3 year grant:

Academic/Industry grant of \$450,000 from Life Sciences Institute (LSI) with industrial partner Chiron.

Received the highest score ever (jointly written with Ph.D student K. Pollard).

Selected to be covered by an article, photos, and video, on the web-site of the Life Science Institute.

Spring 2001: Invited to be Visiting Professor, Department of Biostatistics, Leiden University, the Netherlands, host: Prof. Dr. H. van Houwelingen.

September 2002: **Principal Investigator of NIH Award**, 5 year grant of \$887,664, 2002-2006. Title: “Statistical Analysis of Longitudinal Studies with Gene Expression Data”.

September 2002: **Co-Investigator**, joint UCSF/Berkeley NIH Award, 3 year grant, 2002-2005. Title: “Statistical Analysis of Complex AIDS Cohorts”.

May, 2003: Visiting Professor, Department of Mathematics, Free University Amsterdam.

July 2004: **Principal Investigator of NIH Award**, 2004-2007 (\$1,000,000). Title: “Data Adaptive Estimation in Epidemiology and Genomics”.

June 2004: **2004 Mortimer Spiegelman Award**. The Mortimer Spiegelman Award was established in 1969 by his family and is awarded annually to a young statistician for outstanding contributions in health statistics. It is presented by the Statistics Section of the American Public Health Association (APHA).

September 2004: Selected to be on the cover in portrait-format on one of the five well-respected Tan Applied Mathematics series textbooks, edited by Applied Mathematics for Brooks/Cole, a division of Thomson Higher Education. Quotation from invitation letter: “Famous” applied mathematicians will be featured on the cover of each of the five texts in the hope that seeing a successful applied mathematician will motivate readers (students) of these texts to learn and to use the applied mathematical skills they acquire in their future careers.” Based on this idea, the executive editors of the Tan series have invited me to be featured on one of the five covers of the upcoming new edition.

March 1999 until 2004: Long Term Statistical Consultant at Chiron for the Microarray Technology Research Group, Data Analysis and Method Development. Chiron is the world’s second largest Biotech Company with headquarters located in Bay Area.

April 11, 2005: **2005 van Dantzig Price**.

This is the highest award in Statistics and Decision Theory in the Netherlands.

Once in every 5 years the Dutch Statistical Association presents the Van Dantzig Award to either a dutch statistician or operation researcher under the age of 40.

The award is in memory of prof. dr. D. van Dantzig, the founder of Dutch mathematical statistics.

The former recipients are van Zwet (1970, Statistics), van Meurs (1975, Statistics), Hordijk (1980), Rinnooy Kan (1985), Gill (1990, Statistics), Ridder (1995), and van der Vaart (2000, Statistics).

August, 2005 **2005 Snedecor Award** joint with Nick Jewell.

We received the Snedecor Award for our paper "Case-control current status data" in *Biometrika*, 2004, v91, pp. 529-541.

The criteria for the award are to an individual(s) who has been (1) instrumental in the development of statistical theory in biometry, and (2) who has a noteworthy publication in biometry within three years of the date of the award. So, the award is also a tribute to the overall contribution to biometry.

The award consists of a plaque, a citation, and a cash honorarium. It was presented at the COPSS Awards and Fisher Lecture session at the Joint Statistical Meetings (JSM).

August, 2005 **2005 COPSS Award**.

The Committee of Presidents of Statistical Societies (COPSS) Awards are jointly sponsored by the American Statistical Association, the Institute of Mathematical Statistics, the Biometric Society ENAR, the Biometric Society WNAR, and the Statistics Society of Canada.

The Committee of Presidents of Statistical Societies (COPSS) Award is presented annually to a young member of one of the participating societies of COPSS. The award is presented in recognition of outstanding contributions to the statistics profession. The Presidents' Award is granted to an individual who has not yet reached his or her 41st birthday during the calendar year of the award. The award was established in 1976 and consists of a plaque and a cash award.

September, 2005 **2005 Myrto Lefkopoulou Distinguished Lectureship** at the Biostatistics Department, Harvard School of Public Health.

The lectureship was established in perpetuity in memory of Dr. Myrto Lefkopoulou, a faculty member and graduate of Harvard School of Public Health. Dr. Lefkopoulou tragically died of cancer in 1992 at the age of 34 after a courageous two-year battle. She was deeply beloved by friends, students, and faculty.

Each year the Myrto Lefkopoulou Lectureship is awarded to a promising statistician who has made contributions to either collaborative or methodologic research in the applications of statistical methods to biology or medicine, and/or who has shown excellence in the teaching of biostatistics. Ordinarily, the lectureship is given to a statistician who has earned a doctorate in the last fifteen years. The lecture is presented to a general scientific audience as the first Department colloquium of each academic year. The lectureship includes travel to Boston, a reception following the lecture, and an honorarium. Previous recipients of the Lefkopoulou Memorial Lectureship have been Marie Davidian, Danyu Lin, Bradley P. Carlin, Steven N. Goodman, Ronald Brookmeyer, Michael Boehnke, Trevor Hastie, Hans-Georg Mueller, Giovanni Parmigiani, Kathryn Roeder, and Louise Ryan.

July 1, 2005-2006 **UC Berkeley Chancellor Endowed Chair**.

Spring, 2006 **Miller Professor** funded by the Miller Institute, UC Berkeley.

July 1, 2006- **Jiann-Ping Hsu/Karl E. Peace Endowed Chair in Biostatistics**.

May, 2007 **Charles L. Odoroff Memorial Lecture**, Targeted Learning of Scientific Questions, Distinguished Lecture Award from Department of Biostatistics, University of Rochester.

- July 2007 **NIH-Award Targeted Maximum Likelihood Learning and Super Learning in HIV Research** (2007-2012), 2.3 million dollar grant. Featured by UC Berkeley Sponsored Project Office as Special Award of the week.
- April 22, 2009 The 10-th **Annual Abbott Laboratories Distinguished Lectureship** in Pharmaceutical Applications, A statistics and biostatistics joint seminar, Towards Robust Machine Learning Algorithms for Causal Effects that Preserve Meaningful Statistical Inference, Madison University.
- July 1, 2009 **Distinguished IMS Lecture Award**.
- 2009 RSR Project #09-52 "Clinical Trials: Causal Inference Methodology and its Application in Evaluating Efficacy and Safety of Drugs and Other Medical Products" funded by FDA's CDER's Regulatory Science and Review Enhancement (RSR) Program, involves collaborating with Safety analysis group at FDA.
- 2013 FDA-funded project on HIV safety analysis.
- *2014-2019 **NIH award**, R01AI074345 (PI: van der Laan)
 07/30/2014 6/30/2019, Total: \$1,844,131. Project length: 5 years.
 Targeted Learning: Causal Inference Methods for Implementation Science.
 Major Goal: This project will extend Targeted Learning methodology to address HIV implementation science questions, including hierarchical data, complex dependencies between individuals and clusters, and small sample size. The resulting methods will be implemented as publicly available software, and applied to clinical cohort data from Southern and Eastern Africa and a large cluster randomized trial of antiretroviral-based HIV prevention in Kenya and Uganda.
- *2016-2018 **Bill and Melinda Gates Foundation**, Healthy Birth, Growth, and Development Knowledge Integration Initiative (HBGDki) Targeted Learning Platform. UCB Award ID: 040352-002. PI: Mark J. VAN DER LAAN (UC Berkeley). Total: \$459,757. Project length: 2 years.
- *2016-2019 OPP1165144 (PI: van der Laan), 11/29/2016-12/31/2019
Bill and Melinda Gates Foundation Award, \$2,334,221.
 Title: Causal effects of endogenous and exogenous risk factors for wasting.
 Major Goals: The overall goal is to develop analytics infrastructure to support precision public health and HBGDki (Healthy Birth, Growth and Development knowledge integration) efforts.
- *2016-2018 9223sc (UCB PI: van der Laan), 01/01/2016-05/31/2018
 Bill and Melinda Gates Foundation, \$51,624
 Preterm Birth Initiative
 Major Goals: To contribute to a measurable and sustainable reduction in the burden of preterm birth (PTB). Role: Subaward PI
- *2016-2018 UG3 OD023356 (PI: Eskenazi), 09/21/2016-08/31/2018
 National Institutes of Health, \$908,256
 Long-term Sequelae of Early Life Pesticide Exposure in the CHAMACOS Birth Cohort
 Project Goals: The major goal of this project is to follow more than 600 Mexican American youth in an agricultural community to adulthood for next 7 years, and to monitor effects of early life exposure to pesticides and other pollutants on their growth, health, and learning. Epigenetic and neurodevelopmental pilot studies will

contribute to the development of the master protocols for all ECHO cohorts.
Role: Co-investigator

*2016-2021 R01ES026994 (PI: Eskenazi), 09/30/2016-08/31/2021

NIH, \$704,996

Effect of Early Life Exposure to Social Adversity and Pesticides on Risk-taking Behavior of 16-18 Year Olds: The CHAMACOS Study

Major Goals: To refine methods of assessing exposure to agricultural pesticides and adversity, and determine how these exposures may work together to influence poor behavior outcomes (such as drug use and delinquency) during the critical transition from adolescence to early adulthood.

Role: Co-Investigator

*2013-2018 R01ES007171 (PI: Eskenazi) 08/16/2013-04/30/2018

NIH \$408,785/NCE

In Utero Dioxin Exposure in Seveso, Italy and Health of the Second Generation Major

Goals: In 1976, as a result of an explosion at a chemical plant near Seveso, Italy, the surrounding population was exposed to high levels of 2,3,7,8-tetrachlorodibenzo-p-dioxin or TCDD, a potent chemical that causes cancer and can affect hormones. We are now following the health of the children of the women (Seveso Womens Health Study) who were exposed in Seveso.

Role: Co-Investigator

2014-2018 10185sc (PI: Petersen), 07/01/2017-11/30/2018

NIH, yearly \$186,004

ACTG: Sustainable East Africa Research in Community Health (SEARCH)

Major Goals: The SEARCH study will quantify the health, economic and educational impact of a) early HIV diagnosis and immediate ART (antiretroviral therapy) using a streamlined care delivery system in rural communities in East Africa (Phase I) and b) targeted Pre-Exposure Prophylaxis (PrEP), targeted HIV testing and targeted care interventions on top of universal treatment and streamlined care (Phase II).

Role: Co-Investigator

*07/2017 - 07/2018 Inter-American Development Bank, Evaluating the Performance of the DIABETIMSS Program in Mexico City using Electronic Health Records. UCB Award ID: 043660-001. Co-PI: Mark J. VAN DER LAAN (UC Berkeley). Total: \$113,846. Project length: 1 year.

*2016 **Adrienne Cupples Award** for excellent in teaching, research and service in biostatistics, Boston University.

*2016-2021 **NIH Training Grant Award T32LM012417**

4/1/16-3/31/21, \$1,528,680 (tuition for trainees, and some administration)

Biomedical Big Data Training Program at UC Berkeley

Description: The NIH-funded Biomedical Big Data Training Program at UC Berkeley responds to the urgent need for advances in data science so that the next generation of scientists has the necessary skills for leveraging the unprecedented and ever-increasing quantity and speed of biomedical information. Big data hold the promise for achieving new understandings of the mechanisms of health and disease, revolutionizing the biomedical sciences, making the grand challenge of Precision Medicine a reality, and paving the way for more effective policies and interventions at the community and population levels. These breakthroughs require highly trained researchers who are proficient in biomedical big data science and have advanced skills at collaborating

- effectively across traditional disciplinary boundaries. Biomedical Big Data Training Program at UC Berkeley
- *2017 Small Education grant (3k) for developing online/video lecture/interactive course funding Ph.D student/TA Rachael Philips.
- 2020 ***Novo Nordisk Grant:** 3 million dollar grant over 5 years from Novo Nordisk to my research group on targeted learning for causal inference with real world data.
- 2020 * **200K Accenture Targeted Learning Fellowship:** Funding by Accenture for Accenture Targeted learning Fellowship supporting two Ph.D students, 2 years.
- 2020-2021 * **FDA Demonstration Project of Targeted Learning** funded by FDA, led by Susan Gruber and myself, involving weekly meetings, seminars, monthly meeting with FDA leaders, real world data analysis projects, and three day workshop on Targeted Learning for 265 FDA statisticians.
- 2020- * **FDA funded Innovation Center for Sentinel Project:** led by Pharmacoepidemiology Harvard, one of main tasks validation of Targeted Learning for real world safety analysis.
- 2020-2024: ***NIH RO1 Award on Adaptive Designs:** over 2 million dollar grant.
- 2019- * **Awarded the Turing Lecture**” involving giving a general audience high profile lecture, ”Targeted Learning, the Bridge from Machine Learning to Statistical Inference”, and giving a workshop at the Turing Institute.
- 2022 ** **Research.com ranking #89, #165 among top mathematicians:** Research.com is a leading academic platform for researchers, 2022 Edition of Ranking of Top 1000 Scientists in the field of Mathematics. Ranked #168 in the world ranking and #89 in United States ranking.The ranking is based on the H-index metric provided by Microsoft Academic and includes only leading scientists with an H-index of at least 30 for academic publications made in the area of Mathematics.
- 2019-2022: **Keynote talks:** ASA Biopharm (2019); Turing Lecture (2019); Athlete, Barcelona Institute for Global Health (2020); 3 keynote talks University of Copenhagen, Novo Nordisk, Novo Nordisk Foundation (2020); Philadelphia Innovation Event, organized by Accenture (2020); Causal Inference summit Netflix (2020); 2 keynote talks Novo Nordisk and Novo Nordisk Foundation (2021); Pfizer Analytics Summit (2021).
- 2020, 2022: **ASA Duncan Lecture**, Aug 2020, **Distinguished IMS Lecture** Jan 10 2022.

EDITORIAL WORK AND REVIEW

- ***Associate Editor**, Gastroentology (2016-).
Electronic Journal of Statistics (2007-)
- * **Associate Editor**, Journal of Observational Studies (2014-)
Associate Editor, Statistics Surveys (2005-).
Associate Editor, Lifetime Data Models, 1996-2000.
Associate Editor, Biometrics, 1997-2003.
Associate Editor, Journal of Statistical Planning and Inference, 2001-2006.

* **Associate Editor**, Statistical Applications in Genetics and Molecular Biology, 2002-present.

Associate Editor, Annals of Statistics, November, 2003-2008.

Associate Editor, Founding editor, International Journal of Biostatistics (2004-present).

***Editor 2012-present**: International Journal of Biostatistics.

***Editor, Founding editor, 2012-present** Journal of Causal Inference.

***Associate Editor 2012-present** Epidemiological Methods.

Member of Editorial Board 2006-2008 of ASA-SIAM (Society of Industrial and Applied Mathematics) book series.

Associate Editor, Journal of the American Statistical Association, 2006-2010.

***Associate Editor**, Journal of Statistical Methods in Medical Research, 2006-present.

Co-Editor (with Sandrine Dudoit, and Robert Gentleman) of Special Issue on Genomics of Journal of Multivariate Analysis (2003).

Pharmacoevidence and Drug Safety Best Reviewer, 2014.

Proofread an introductory book in Statistics of Prof. Dr. E. Lehmann (1997).

Book chapter review of book on adaptive designs (2014).

Book review (1997), "Probabilistic Causality in Longitudinal Studies," for Statistics in Medicine.

Book review (1997), "Problems and Solutions in Biostatistical Theory", Duxbury Press, Brooks/Cole Publishing Company.

Book review (1998), Introduction in Biostatistics, Text book, Duxbury Press, Brooks/Cole Publishing Company.

Book review (1999), Statistical Consulting: A Guide to Effective Communication, by Janice Derr, Duxbury Press, Brooks/Cole Publishing Company.

Book review (1999), Fundamentals of Biostatistics, Fifth Edition, by B. Rossner.

Book review (2004), Introduction in Computational Biology.

Book review (2004), The False Discovery Rate, by Yoav Benjamini, Cambridge University Press.

Reviews of NSF Research Proposals (1998, 2 in 1999, 2 in 2000).

Review of NSF Research Proposals, Washington, October 17-19, 2002.

Review of research proposal for Center in Genetic Epidemiology, University of California, Irvine, (2002).

Review of Research proposals of the National Science Foundation of the Netherlands: (2003).

* Review of NIH grant, 2006-

Review grant proposal Inter-American-Development Bank (IDB) (2018).

Review of PCORI grants 2012.

Review of NSF the Netherlands (2010), (2011), (2012), (2013), (2014),(2016), (2018);

Review of Discovery grant proposals, Mathematics and Statistics Evaluation Group (2014).

Reviewing research proposal for Chancellor office UC Berkeley (2013, 2020).

* regularly review articles for journals I am not an AE for such as Annals of Statistics, JASA, Epidemiology, Statistics in Medicine, etc.

* (2021) Reviews of Vici-program grant proposals, most competitive grant applications only awarded to the very best in the Netherlands.

CONSULTING POSITIONS

- 1995: Cost-effectiveness of physicians in San Francisco Hospitals for M.D. I. Ahwah: What variables of a patient in an emergency room predict sensible cost best?
- 1996: Cost-effectiveness of physicians in San Francisco Hospitals for M.D. I. Ahwah: Relation between acuity of a patient and sensible utilization of items by revenue center.
- 1996: Consultant on studies of the influence of caffeine on birth defects for Laura Fenster, California State Department of Health.
- 1998: Consultant on NIH grant “A Nonparametric MLE Survival Analysis Module” with P.I. Dr. Y. Zhan, Data Analysis and Products Division, Mathsoft, Inc.
- 1997: Statistical analysis of relation between monthly budget patient days and monthly actual patient days, Children’s hospital, Oakland.
- 1998: Member of consulting group, consisting of faculty members of the Statistics Department, for statistical problems presented by NSA.
- 1998: Consultant on Cystic Fibrosis Foundation Project “Effects of Flavonoids on Nasal PD in Cystic Fibrosis Patients”, Pediatric Clinical Research Center, Children’s Hospital, Oakland. Part of this project is concerned with linking the genotype of the Cystic Fibrosis gene to the clinical parameters.
- 1999-2002: Statistical consultant for CHIRON for analyzing gene expression data; coordination of statistical analysis in the Microarray Research Group at Chiron.
- 2005-2009: Statistical consultant for the Buck Aging Institute, Novato, CA, for the Bioinformatics core.
- 2007-: Statistical consultant AmGen.
- January 2008-2012: Founder and CEO of Target Analytics.
- 2008-: Statistical consultant for pharmaceutical company.
- 2008-: Statistical consultant for legal cases.
- 2011-: Statistical consultant for Worldbank.
- January 2010-2012 : Statistical consultant for world bank, implementation science, design of observational studies.
- 2010-2015: Statistical consultant FDA safety analysis group.
- 2012: Statistical consultant for Jansen Jansen on observational data analysis for treating Alzheimer.
- *2011-: Statistical consultant for Kaiser Permanente
- 2012-2018: Statistical consultant for Metronomx. Developed design and analysis plan, and successfully presented at FDA for drug-approval.
- 2014: Tax assessment consulting, Winston & Shawn.
- 2014: Statistical consultant Department of Health Care policy, Harvard University, development of dynamic treatment strategies for depression based on data from VA.
- 2016-2018: Statistical consultant Nielsen Inc.
- 2017: consulted on legal patent infringement case.
- 2016-2018: Member of advisory board of FDA grant of Michael Rosenblum.

- 2016-2018: Statistical consultant for Opioid PMR Consortium (OPC) Observational Studies Workgroup (OSW), FDA guided observational study to evaluate effect of medium and longlasting opioid medications on death/overdosis/addiction.
- 2020-2022: *Statistical consultant Accenture, San Francisco.
- 2021-: *Statistical consultant Parexel.
- 2019-2022: *Statistical consultant FDA funded Targeted Learning demonstration projects, including one in the Sentinel Project.
- 2021: *Evaluation of expert witness in legal case, deposition based on this as well.

University Service (2018-)

1. Chair Ad Hoc committee for promotion of SPH faculty member to Full Professor, 2022.
2. Chair mid career review (2022).
3. Chair Ad Hoc committee promotion of SPH faculty member to Full Professor 2021.
4. Campus Ad Hoc Review Committee for tenure case (2020).
5. Member Academic Personnel Committee 2020-2022.
6. Member admissions committee for MA and Ph.D Statistics Department (2020-2021).
7. Admissions committee for MA and Ph.D Biostatistics, each year.
8. Co-Chair of group of Biostatistics (2018-).
9. Academic Director Center of Targeted Learning (2016-)
10. Meetings with important visitors of UCB: e.g., met with former CEO Research Merck, Perlmutter, this year when visiting Robert Tijan, among others.
11. Member of the search committee for a biostatistics assistant-professor (2021-2022).
12. Assist search and interviewed faculty for the new Computational Precision Medicine Department led by Maya Petersen.
13. I am a member of the computational biology group on campus;
14. I am a member of admissions committee for the Designated Emphasis in Computational Biology;
15. I am co-Director of the Biostatistics and Computing core of the Superfund Research Program (Genomics) in the School of Public Health headed by Prof Martyn Smith (2004-);
16. I am formal supervisor of our IT person B. Bundy and statistical programmer J. Schwab.
17. I am actively engaged in recruiting students for our MA and Ph.D program.
18. Salary equity analysis for SPH with Prof Hubbard 2020.

INVITED LECTURES

1. *Efficient Estimation in Nonparametric Missing Data Models*: September 17, 1992, EMS Conference, Bath.

2. *Hoffmann-Jorgensen Weak Convergence Theory and the Proof of an Almost Sure Representation Theorem*: March 20, 1991, CWI, Amsterdam. Workshop *Statistics in Large Parameter Spaces*.
3. *General Efficiency Theory for the NPMLE and an Identity for Linear Parameters in Convex Models*: April 15, 1993, Euler International Mathematical Institute, St. Petersburg, Russia. Workshop on *Nonparametric and Semiparametric Models (asymptotic problems)* of the Kolmogorov Semester on Probability and Statistics. *The Bivariate Censoring Model*: April 11, 1994, Cleveland, Ohio. Invited Speaker at the 1994 Biometric Society ENAR spring meeting on the subject "Multivariate Censored Data", held jointly with the IMS and ASA.
4. *Proving Efficiency in Biased Sampling and Missing Data Models*: June 24, 1994, Chapel Hill, North Carolina. Invited Speaker for the session "Likelihood" at the 3rd World Congress of the Bernoulli Society and 57th Annual Meeting of the Institute of Mathematical Statistics. *Locally Efficient Estimation with High Dimensional Covariate Processes*, December 16, 1994, Oberwolfach, Germany. Invited speaker for Conference on "Asymptotic Methods for High Dimensional Data". *Singly and Doubly Censored Current Status Data: Estimation, Regression and Asymptotics*, August 2, 1995, Berkeley. Invited speaker, NSF Econometrics Symposium 1995.
5. *An Identity for NPMLE in Censored Data Models*, June 24, 1996, Washington State University, Pullman, WA. Invited speaker, Joint Regional Meeting Biometric and IMS.
6. *Locally Efficient Estimation with Current Status Data and Covariates*, August 5, 1996, Chicago. Invited speaker, Joint Statistical Meeting.
7. *Nonparametric Estimation of the Bivariate Survival Function*, August 26, 1996, Vienna, Austria. Invited speaker, 4th World Congress of the Bernoulli Society.
8. *Inference in High Dimensional Semiparametric Censored Data Models*, July 7-9, 1997, Taipei, Taiwan. Speaker and organizer of session, Joint meeting with the Chinese Statistical Association and the Chinese Institute of Probability and Statistics.
9. *Inference in High Dimensional Semiparametric Censored Data Models and Testing Treatment Effects in Observational Studies*, August 18-22, 1997, University of Minnesota, Minneapolis, Minnesota. Invited speaker, Workshop of Institute of Mathematics and Applications.
10. October 1997, Invited by Prof. Dr. J.M. Robins, Harvard School of Public Health, Boston, to give series of lectures on proving asymptotics for semiparametric models.
11. Fall, 1997, Invited speaker, Statistics Seminar, Emory University, Atlanta (also invited at another University in Atlanta, but cancelled).
12. March 3, 1998 Invited speaker, Statistics Seminar, Department of Statistics, University of California, Berkeley.
13. September, 1998, Speaker Biostatistics Seminar, UC Berkeley.
14. *Locally Efficient Estimation in Censored Data Models: Theory and Examples*, October 22, 1998, Invited speaker, Statistics Seminar, Department of Statistics, Florida State University, Tallahassee.
15. *Causality in Public Health Studies*, October 29, 1998, Invited speaker, Foundations for the Future Symposium in honor of Dean P. Buffler, School of Public Health.

16. *Locally Efficient Estimation in Censored Data Models: Theory and Examples* October 23-25, 1998, Invited speaker, Conference in honor of Professor Alfred H. Clifford, Tulane University, New Orleans, Texas. Clifford lecturer: Bickel.
17. *Current Status Data on a Stochastic Process*, August 8-12, 1999, Invited speaker and discussant Joint Statistical Meeting ASA, Baltimore.
18. *Locally Efficient Estimation with Multivariate Right Censored Data*, May 15-20, 2000, Invited speaker Survival Analysis session of the 5-th World Congress of the Bernoulli Society for Probability and Mathematical Statistics, Guanajuato, Mexico.
19. *Statistical Inference with Microarray Data using the Parametric Bootstrap*, August 25, 2000, National Cancer Institute, Washington.
20. *Statistical Inference with Microarray Data using the Parametric Bootstrap*, March 28, 2001, Empirical Processes in Biostatistics, invited session ENAR/IMS.
21. Invited speaker of the SCI 2002 Sixth Multi-Conference on Systemics, Cybernetics and Informatics, July 14-18 Florida.
22. Invited speaker IISA Fourth Biennial International Conference on Statistics, Probability and Related Areas, June 14-16, 2002.
23. Invited speaker Genomics Seminar, University of California, Berkeley, January, 2003.
24. Invited speaker at a Symposium "Challenges in the Statistical Analysis of Genomic Data" at the AAAS meeting, Denver, February 13-18, 2003.
25. Invited speaker in session on genomic data of the Joint Statistical Meeting, August 3-7, 2003.
26. Two lectures as the Constance van Eeden Visiting Professor, University of British Columbia, Vancouver, Canada, October 6-12, 2003.
27. Invited speaker, International Conference on Analysis of Genomic Data, the Harvard Medical School, May 10-11, 2004, Boston.
28. Invited Keynote speaker, Taipei Symposium on Statistical Genomics, Institute of Statistical Science, Academia Sinica, December 15-17, 2004.
29. Invited Odoroff Memorial lecture on Targeted Maximum Likelihood Learning, Department of Biostatistics, University of Rochester, September 20, 2007.
30. Invited Lecture on Causal Inference in Clinical Trials and Post Market Data Analysis, Federal Drug Administration (FDA), August 8, 2007.
31. Invited Miller Lecture on Super Learning, Miller Institute, as Miller Professor, May 3, 2007.
32. Invited lecture on Targeted Maximum Likelihood Learning and Super Learning in AIDS Research, April 11-13, Workshop on statistical methods in AIDS research, Vaile, Colorado.
33. Invited lecture on Targeted Maximum Likelihood Learning, ENAR, 12007, April 14 (lecture given by my Ph.D student Dan Rubin).
34. Invited presentation and organizer of Adaptive Designs Session at International Statistical Institute Meeting, Durham, South Africa, August 16-22, 2009.
35. Invited Distinguished Lecture, IMS meeting in Seoul, June 28-July 1, 2009.
36. Prominent researchers special invited session on Causal Inference in High Dimensional Applications, Joint Statistical Meeting, Washington, August 2-7, 2009.

37. Invited speaker, Statistics Seminar, Leiden University, the Netherlands, July 2009.
38. Invited speaker, Webinar (Amstat, biopharmaceutical session), attended by hundreds of industry statisticians, August 24, 2009. Presentation on genomic data analysis and biomarkers.
39. Organizer of Session on Recent Advances on Adaptive Designs at 57-th ISI Conference, Durban, South Africa, August 2009, Michael Rosenblum gave presentation on this session on our joint work, other speakers Scott Emerson and Bruce Turnbull.
40. Workshop on Causal Inference for the FDA statisticians at the FDA, September 19, 2009.
41. Invited speaker in workshop discussing statistical methods for causal effect assessment in safety analysis in HIV, in particular, relation between abacavir and cardiovascular disease, Amsterdam, May 10-12, 2010.
42. Invited speaker on adaptive designs in HIV prevention trials, June 2, 2010, Bill and Melinda Gates foundation, Seattle.
43. Invited speaker two seminars in the Netherlands on causal inference, June 10-17, 2010.
44. Invited discussant of session on dynamic treatment regimens, JSM, August 2010.
45. Invited speaker of workshop on community based interventions, Pefpar and Bill and Melinda Gates Foundation, Washington DC, October 6-8, 2010.
46. Invited speaker to give 4 lectures on Targeted Maximum Likelihood Estimation, Super Learning, and Causal Inference in 42 Winter Conference, Hemavan, Sweden, March 6-9, 2011.
47. Invited speaker, French Berkeley Research collaboration, Paris Descartes, January 2011.
48. Invited speaker, Statistics Colloquium, which is held jointly by four Montreal universities: McGill, Université Montreal, Concordia and Université du Québec à Montreal, March 25, 2011.
49. Invited to Division of AIDS-NIAID sponsored workshop on Quantitative Methods to Advance the Combination HIV Prevention, Gates Foundation, September 6-7, 2011.
50. ICSA, invited lecture in session Causal Inference and its applications in drug development , June 28, 2011.
51. Galician Society for Statistics and OR, Plenary Talk during the X Galician Conference (Pontevedra, Spain; November 2-5 2011).
52. Invited speaker, CRM Causal Inference in Health Research workshop, Montréal, May 9-13, 2011.
53. Invited lecture on observational longitudinal studies and causal inference, August 17-19, 2011, ISI satellite meeting on Dynamic Statistical Models, Copenhagen, Denmark.
54. Invited talk at Advanced Topics in Pharmacoepidemiology (114) 28th International Conference on Pharmacoepidemiology & Therapeutic Risk Management, Barcelona, August 22-26, 2012.
55. TMLE based approach to confounder selection, Invited talk at the International Biometric Conference, Kobe, Japan, August 26-31.

56. Invited speaker on comparative effectiveness research, IMS-China 2011, Xian, China, July.
57. Invited speaker, ENAR 2012, "Multivariate statistics in high-dimensional data", Washington DC.
58. Presentation at FDA to present novel statistical analysis plan for approval of drug for Chagas disease, December 16, 2012, Washington DC, part of consulting for Metronomx.
59. Invited speaker ENAR March 11, 2013, Orlando, Florida, new developments in the construction and optimization of dynamic treatment regimes.
60. Invited speaker, CMS Winter Meeting 2012 in Montréal, Targeted Learning in HIV.
61. Invited speaker, 7th Conference of the EMR-IBS to be held in Tel Aviv, Israel on 22-25 April 2013, and to speak in the session entitled "Graphical models/machine learning methods in biostatistics" .
62. Annual meeting Statistical Society of Canada, Edmonton, 2013, May 29.
63. Invited lecture, LeDell, Erin; Petersen, Maya L.; and van der Laan, Mark J. Computationally Efficient Confidence Intervals for Cross-validated AUC Estimates. Joint Statistical Meetings, August 2013. Montreal, Canada.
64. ENAR March 2014, Invited speaker session on networks.
65. Invited speaker Johns Hopkins University, Targeted Learning of Optimal Dynamic Treatment, Biostatistics, March 2014.
66. Invited speaker Colloquium Foundation of Statistics, revisiting foundations for era of Big Data, February 21, 2014.
67. Invited speaker, Targeted Learning of optimal dynamic treatment, Department of Pharmacoepidemiology, Harvard University, February 18, 2014.
68. Invited speaker Causal Inference seminar, School of Public Health, Harvard University. Targeted learning of optimal dynamic treatment and networks. February 19, 2014.
69. Invited speaker, Annual Meeting of the French Statistical Society, Rennes, France, June 2-6, 2014.
70. JSM 2014, invited to present award to Judea Pearl.
71. Special invited speaker, Targeted Learning for Optimal Dynamic Treatments ISIS conference, July 10-13, 2014.
72. Discussant for session on recent advances in causal inference, August, JSM 2014, Boston,
73. Lecture in workshop Data Driven Discovery, Moore Foundation, one of 28 final candidates for 14 grants awarded by Moore Foundation, July 28-29, 2014.
74. Invited talk, Targeted Learning in Precision Medicine, Simons Institute, Feb, 2015.
75. Invited talk, JSM 2015, Health Policy Statistics Section, Collaborative Targeted Maximum Likelihood Estimator (CTMLE) in Observational Studies. Seattle, August 2015.
76. Invited talk at FDA, Targeted Learning, September 2015.

77. Invited talk, Targeted learning of the optimal dynamic treatment and Statistical inference for the mean outcome under the optimal dynamic treatment, session Improving Medical Decision Making in the Era of Personalized Medicine, 11th International Conference on Health Policy Statistics taking place October 7-9, 2015 in Providence, Rhode Island.
78. Invited talk, Online targeted learning, The National Institute of Allergy and Infectious Disease (NIAID), Biostatistics Research Branch (BRB), Infectious Disease Research Quantitative Methods and Models in the Era of Big Data Statistical Workshop, November 9-10, 2015 in Bethesda, MD.
79. Invited talk, session is for the 4th IMS APRM (Institute of Mathematical Statistics Asia Pacific Rim Meeting), June 27-30, 2016, Hong Kong.
80. Invited talk (wide audience), Targeted Learning, Information, Operations and Management Sciences colloquium at the NYU Stern School of Business, October 2016.
81. Invited seminar, The KP Center for Effectiveness & Safety Research (CESR), Targeted Learning: Applications in Precision Medicine, January 14, 2016.
82. Invited talk, Highly Adaptive Lasso, Atlantic Causal Inference Conference, New York, May 2016.
83. Invited discussant, Super-learner versus sensitivity analysis, ACIC, New York, May 26-27, 2016.
84. Invited talk, IBC 2016, Generally Efficient TMLE for Arbitrary Models and Target Parameters: A Super-learner guaranteed to converge at faster rate than $n^{-1/4}$. July 2016, Victoria, Canada.
85. Invited talk, workshop on High-Dimensional Causal Inference and its Application to Genetics, Centre de Recherche Mathématique, Montreal, July 25th-29th 2016.
86. Invited award lecture, Targeted Learning in Precision Medicine, Boston University, 2016 Adrienne Cupples Award lecture, April 7, 2016.
87. Invited talk, Targeted Learning of Optimal Subgroups in Precision Medicine, ENAR March 6-9, 2016, Austin.
88. Invited talk, JSM 2016, Targeted Learning in Precision Medicine, August 2, 2016.
89. Invited discussant, JSM 2016, of JASA-T&M Special Invited Paper Personalized Dose Finding Using Outcome Weighted Learning, Kosorok et al. , Chicago, July 30 - August 4, 2016.
90. Invited talk, Targeted Learning of Causal Effects of Interventions on a Survival outcome, seminar within International Scientific Meeting on Survival Analysis of Population-based data, August 31-September 2, 2016, London School of Hygiene and Tropical Medicine.
91. Invited talk, Targeted Learning in Precision Medicine, Workshop at UC Berkeley, Global Alliance with Cambridge University and Singapore University, March 28-29, 2016.
92. Invited seminar, Kaiser Permanente Big Data Seminar, Targeted Learning in Precision Medicine, January 14, 2016.
93. Taught workshop on Targeted Learning, jointly with Susan Gruber, 2nd Seattle Symposium on Healthcare Data Analytics, October 23, 2016.

94. Targeted Learning: Causal Inference in Observational and Experimental Studies. Presentation at visit of Calico company, including CEO Art Levinson, January 18, 2017.
95. Invited presentation, ENAR 2016, Optimal subgroups, March 6-9, 2016, Austin, TX.
96. Invited talk, Causal inference for network data, workshop on Causal Inference with Highly Dependent Data in Communicable Diseases Research, Harvard University, May 25, 2016.
97. Targeted Learning, seminar statistics department, UC Berkeley, October, 2016.
98. Invited talk, Targeted Learning, Liver forum meeting in Barcelona, April 12, 2016.
99. Taught Workshop on Big Data for personnel of the Instituto Mexicano del Seguro Social, jointly with Alan Hubbard and Wenjing Zheng, November 3-4 2016, Mexico City, Mexico.
100. Invited talk, Targeted Learning, Biostatistics Seminar at Penn, 11/01/2016.
101. Invited talk, Fred Hutchinson Cancer Center, One-Step Targeted MLE and Highly Adaptive Lasso, October 19, 2016.
102. Invited talk, Causal Inference Conference 2016 , One-Step Targeted Maximum Likelihood Estimation and the Highly Adaptive Lasso SQM at Michigan University, November 9, 2016.
103. Invited talk, Environmental Knowledgebase Workshop, BIDS, UCB, January 31, 2017.
104. Invited talk, Yale University, Quantitative Research Methods Workshop, February 2, 2017.
105. Keynote talk, VVS (National meeting Dutch Statistical Association), Targeted Learning in Healthcare, Theme Healthcare for the future, March 23, 2017, Utrecht, the Netherlands.
106. Invited panel presenter, Causal inference, Computational Health Science symposium-UCB/UCSF/LBNL, April 25, 2017, UCSF.
107. Taught workshop on Causal Inference and Adaptive Design, jointly with Maya Petersen, HIV/Liver Forum, Washington DC, May 10, 2017.
108. Invited discussant, ACIC 2017, Interference and Social Network, University of North Carolina at Chapel Hill, May 23-25, 2017.
109. Invited talk, Asymptotically efficient estimation under no assumptions, Session title All models are wrong, SER, June 21-24, 2017, Seattle.
110. Invited special session on "Data representation with graphs and multi-scale processing" organized within the next conference Wavelets and Sparsity XVII, San Diego Convention Center, San Diego, California, August 6 -10, 2017.
111. Invited talk, Targeted Learning, Department of Psychology, advertised broadly, University of Groningen, September 11, 2017.
112. Invited talk, Targeted Learning, workshop in honor of Richard Gill, Statistical theory and the real world, Lorentz Center, Leiden University, September 12-15, 2017.
113. Taught workshop on Longitudinal Causal Inference, jointly with Sherri Rose, Columbia University, September 22-23. 2017

114. Invited talk, Colloquium series of the Statistics, Department of Statistics, Penn State, October 12, 2017.
115. Keynote talk, Targeted Learning and Causal Inference, workshop statistical methods in epidemiology, Braunnshweig, Germany, November 23-24, 2018.
116. Invited guest lecture, HAL-TMLE, Causal Inference Seminar Series at UMass, November 29, 2017.
117. Invited talk, Targeted Machine Learning and Statistical Inference, Workshop on the Interface of Machine Learning and Statistical Inference, Banff International Research Station (BIRS), January 14-20, 2018.
118. Invited talk, Targeted Learning in Precision Medicine, ENAR 2018, Invited session on Precision Medicine and Artificial Intelligence, March 25-28, 2018, Atlanta.
119. Invited presentation Targeted Machine Learning and Adaptive designs, at the Research Showcase April 21st, 2018, School of Public Health, Cal Day.
120. Invited visitor, Antoine chambaz, Department of Mathematics, Paris Descartes, May 14-21, 2018.
121. Invited talk, second International NASH Biomarker workshop, May 18-19, 2018, Washington DC.
122. Invited presentation, Targeted Machine Learning for Causal Inference. ACIC 2018, Causal Inference with High-Dimensional Covariates, May 22 and 23, 2018.
123. Invited discussant, Collaborative TMLE, Ten Have & Biometrics Invited Session at ACIC 2018, May 23, 2018, Pittsburg.
124. Invited panel discussant, National Academy of Science, Workshop - Examining the Impact of Real-World Evidence on Medical Product Development: A Three-Part Workshop Series, July 17-18, 2018.
125. Invited presentation, JSM 2018, Causal inference meets statistical learning with complex data, August 2018, Vancouver.
126. *Invited presentation, Optimal Dynamic Treatments and Randomized Trials, Department of Radiation Oncology, UCSF, August 27, 2018.
127. Invited Lecture Causal inference group, UNC, October 5, 2018.
128. Giving 1-day workshop on Targeted Learning, jointly with Susan Gruber, 2018 Seattle Symposium on Health Care Data Analytics, October 2018.
129. *Invited speaker, causal inference session for survival data, Lifetime Data Science Conference, Pittsburgh, May 29-June 1, 2019.
130. *Invited talk, Workshop Leiden University, Celebrating Aad van der Vaart's contributions to statistics and 60-th birthday. June 17-21, 2019.
131. *Invited talk, 2019 Joint Statistical Meeting, Extraordinary opportunities for utilizing real-world evidence to impact drug development and regulatory decision-making, August 2019.
132. * Keynote presentation on Targeted Learning, ASA Biopharm, Washington DC, September 24, 2019.
133. * Invited seminar presentation, Department of Biostatistics, Columbia University, October 17.

134. Invited workshop on targeted learning, Annual Deming Conference, December 4-6, 2019, Atlantic City.
135. *Invited talk, SAMSI causal inference workshop, December 9-11, 2019.
136. * April 8-9, 2020, Invited speaker, Data science workshop jointly sponsored by NHLBI and NLM NIH main campus. Invited speaker.
137. * Keynote Turing Lecture at Turing Institute, combined with workshop on Targeted Learning the previous day, London, March 2-5, 2020.
138. * February 6, 2020, Invited talk, Princeton Seminar Quantitative Social Science Colloquium, Princeton.
139. * Keynote speaker at Kick-off consortium meeting, Athlete, exposome from Evidence to translation, Barcelona Institute for Global Health, Barcelona, January 27-28, 2020. Introduction of super-learning and targeted learning as general statistical tools in environmental science.
140. * Invited talk on Targeted Learning in causal inference seminar, Department of Statistics, UC Berkeley, Fall 2019.
141. * Workshop on Targeted Learning, Accenture, SFO, January 14, 2020.
142. * FDA webinar about Targeted Learning for 100 FDA epidemiologists and statisticians, February 5, 2020.
143. * Workshop on Targeted Learning, CSP, Sacramento, February 20, 2020.
144. * Invited talk Econometrics seminar, UC San Diego, May 2020.
145. * Invited keynote talks, University of Copenhagen and Novo Nordisk, March 10-15, 2020.
146. * May 19-21, 2020, Invited talk, TMLE of Causal Impacts of Multiple Time-Point Interventions on Survival, Lifetime Data Science Conference, Raleigh, NC.
147. * August 12, 2020, Invited talk, ASA Duncan Lecture, Mathematical Science Institute.
148. * Invited talks at JSM 2020, Discussant Latent variables. Targeted Learning of Causal Effects on time to event.
149. * August 12, 2020, Adaptive Designs for Epidemic Control: Test and Respond Strategies for COVID-19, MSRI workshop on Epidemics.
150. * August 2020, Invited talk for AISC machine learning community, University of Rochester.
151. * September 2020, Keynote speaker, Philadelphia Innovation Event, organized by Accenture.
152. * Invited talk at ICPE symposium October 7, 2020, Machine Learning to optimize confounding control based on studies using high dimensional health care data.
153. * Invited keynote speaker, Causal Inference Summit, Netflix, October 30, 2020.
154. * September 23, 2020, Sentinel Innovation Center seminar, Targeted Learning - The Bridge from Machine Learning to Statistical and Causal Inference.
155. * December 14, 2020, Yearly Edwards Lifesciences Pocock Seminar, discussant. December 2020.
156. * January 19, 2021, Invited talk for Online causal inference seminar, Stanford, Higher order TMLE with applications to causal inference, with discussant Alex Luedtke.

157. * March 4, 2021, Invited seminar Buffalo University.
158. * May 11, 2021, Q and A at Novartis with Martin Ho from FDA, Causal Inference Frameworks for Study Design and Analysis of RWD.
159. * June 18, 2021, Invited speaker, panelist, SMAC 2021, Statistics, Philosophy and Health, Virtual.
160. * May 28, 2021, Invited speaker Frontiers of causal inference in data science: Perspectives from leaders in tech and academia. over 1500 people registered.
161. * March 22-24, 2021, Survival Prediction Symposium, AAAI symposium, Invited speaker, Targeted Maximum Likelihood Estimation of the Causal Effect of Treatment on Survival.
162. * March 14-17, 2021, ENAR 2021, short course and invited presentation Targeted Learning for causal inference based on real world data.
163. * September 2021, Invited Causal Inference Seminar at Stanford, Higher order TMLE.
164. * July 18-20, 2021, Invited speaker, 42nd Conference of the International Society for Clinical Biostatistics (ISCB 2021), to be held virtually.
165. * September 11-12, 2021, Invited speaker, Causal Inference Conference, Peking University, virtual, higher order TMLE with applications to causal inference.
166. * August 8-10, 2021, Invited discussant for session on latent variables, and invited panel on assessing PhD readiness in Biostatistics, JSM 2021.
167. * April 3, 2021, Invited seminar Causal Inference from electronic health records, Deep Medicine Research Group, Oxford, Targeted Learning: the bridge from machine learning to statistical and causal inference.
168. * April 23, 2021, Invited speaker, Targeted Learning based on real world data, 19-th ASA Annual connecticut chapter mini conference: borrowing from historical and real world data: latest methodology development and applications,
169. * September 23, 2021, Invited talk, seminar University of Michigan, Ann Arbor, MI.
170. * October 25-27, 2021, Invited workshop on Targeted Learning, BASS, Charlotte NC.
171. * October 11-12, 2021, Invited talk, Novo Nordisk collaboration workshop, UC Berkeley, University of Copenhagen, Oxford University, Oxford, United Kingdom.
172. * October 13, 2021, Invited keynote talk, Novo Nordisk A/S Symposium.
173. * October 14, 2021, Invited keynote talk, Understanding cause and effect through data science and novel biomedical data sources, Towards integration of targeted learning with real world data into the drug development process, Novo Nordisk Foundation, Copenhagen, Denmark.
174. * Invited for workshop June 14-18, 2021, Statistics, Philosophy and Health.
175. * June 7, 2021, SER, Workshop - Targeted Learning: Causal Inference Meets Machine Learning.
176. * September 2, 2021, Invited speaker, Higher order TMLE, Seminar University of North Carolina, Chapel Hill, NC.
177. * Invited speaker workshop on Causality in sequential decision making” workshop at NeurIPS, December 6-14, 2021.

178. * Keynote speaker Pfizer Analytics Summit, which is held annually at Pfizer. The occasion provides a forum for the Pfizer analytics community to share best practices in the use of modern machine learning and AI tools in drug development. September 21-22, 2021.
179. * September 16, 2021, Invited speaker, FDA Webinar. Targeted Learning: towards a future informed by real world evidence.
180. * September 21-24, 2021, Invited speaker, ASA Biopharmaceutical section regulatory-industry statistics workshop, BIOPHARM 2021, Session Title: Key Steps in Generating Real-World Evidence from Analysis of Real-World Data.
181. * November 8, 2021, Invited speaker on webinar Statistical Analysis and Data Quality, Real world evidence, organized by Liver Forum.
182. * December 18-20, 2021, Invited talk on Higher order TMLE, Hybrid CFA-DMStatistics 2021, London
183. * January 10, 2022, workshop on Statistical Methods in Genetic/Genomic Studies, Distinguished Lecture, Targeted Learning with applications to genomic studies.
184. * January 24-28, 2022, Winter School Machine Learning, Biostatistics and artificial intelligence in modern biomedical research, Targeted Learning with Applications to Causal Inference, University of Padova, Italy.
185. * March 14-18, 2022, Workshop on Targeted Learning for 265 FDA statisticians, joint with Hana Lee, Susan Gruber, Rachael Phillips, part of FDA funded demonstration project on Targeted Learning.
186. * March 29, 2022, invited speaker in session on data driven discovery and decision making, ENAR 2022, Highly Adaptive Lasso.
187. * May 17, 2022, Invited speaker session on When and how to use covariate adjustment to improve precision in randomized trials, 43-rd SCT Annual Meeting, Society for Clinical Trials meeting, San Diego, CA.
188. * May 23, 2022, workshop on Targeted Machine Learning of the Causal Effects of Dynamic and Shift Interventions with the tIverse R Packages, Atlantic Causal Inference conference, Berkeley, CA. Co-organizer of ACI 2022.
189. * June 9-11, 2022, First International Conference on Causal Interactive Learning, Washington DC, Higher order Targeted Maximum Likelihood Estimation.
190. * June 22, 2022, Invited speaker, Targeted Learning for Generating Real World Evidence, DIA Global Annual Meeting, Chicago.
191. * June 6, 2022, Invited speaker Symposium on Risks and Opportunities of AI in Clinical Drug Development, Roux Institute at North Eastern University, Towards Integration of Targeted Learning and Causal Inference in Drug Approval Process and Safety Analysis
192. * August 8, 2022, Joint Statistical Meeting, Washington DC. Invited panel, Machine Learning and Artificial Intelligence: Uses and Misuses sponsored by ASA section for statistical programmers and analysts.
193. * August 11, 2022, Joint Statistical Meeting, Washington DC. Invited speaker Recent developments in causal inference with real world evidence in drug development. Targeted Learning of Causal Effects with Continuous Time-to-Event Outcomes.

194. * August 21-26, 2022, Statistical Challenges in the identification, validation, and use of surrogate markers. Invited speaker, "The oracle surrogate and sequential adaptive designs that learn optimal individualized treatment rules by utilizing surrogate outcomes", Oaxaca, Mexico.
195. * September 6-7, 2022, JICI workshop, Berkeley, CA. Organizer and speaker on Federated Learning.
196. * November 3-4, 2022, Invited speaker in session on Real World Evidence, Bay area Biotech-pharma Statistics Workshop (BBSW), Foster City, CA.
197. * June 11-14, 2023, Invites speaker in Session on machine learning methods with biomedical related applications, University of Michigan, Ann Arbor. Targeted Learning and Causal Inference for Integrating Real World Evidence into the Drug Approval Process and Safety Analysis.
198. * March 19, 2023, Short course on "Targeted Learning: Advanced methods for causal machine learning", ENAR, Nashville.

TEACHING

University of California, Berkeley Statistics, Public Health (Biostatistics), and Bioengineering departments

- STAT 102 Introduction to Theoretical Statistics; Spring 1994
- PH 142A Introduction to Probability & Statistics in Biology and Public Health; Fall 1994, 1995, 1996
- PH 142AB Introduction to Probability & Statistics in Biology and Public Health, Summer course; Summer 1996
- BE 190C Statistical methods for clustering, regression and prediction as part of this course on Computational Biology for undergraduates; Fall 2002
- STAT 210B Theoretical Statistics for Ph.D students; Spring 2005
- PH 240A Introduction to modern biostatistical theory and practice; Spring 2015, 2016, 2018, * 2021, *2022.
- PH 240B/Stat C245B Biostatistical Methods: Survival Analysis; Spring 1997, 2000, 2002, 2004, 2006, 2008, 2010, 2012
Spring 2014, Spring 2016, Fall 2017, *Fall 2019, *Fall 2020, *Fall 2022
- PH 243B Special Topics in Biostatistics: Asymptotic Methods in Statistics, Spring 1995; Censored Data and Regression, Spring 1996, Spring 2001; Causal Inference, Spring 1999, Fall 2002; Statistical Techniques in Computational Biology, Fall 2001; Multivariate Statistical Methods in Genomics: Multiple Testing and Loss Function Based Estimation, Fall 2003, 2005, 2007, 2009; Adaptive Designs and Targeted Maximum Likelihood, Spring 2003, 2005, 2007, 2009.
- PH 243D Adaptive designs; Fall 2010, 2012, 2018.
Fall 2018.
- PH 246A/STAT C249A Censored Longitudinal Data and Causality; Fall 2011, Spring 2013
- PH 252B Causal Inference in Longitudinal Studies; Fall 2004, 2006, 2008, 2010, Spring 2013

* PH 295, Targeted Learning with Biomedical Big Data; Fall 2016, Spring 2018, *Spring 2019, *Fall 2021, *Fall 2022.

Instructional Strategy: Most pedagogical studies (i.e., those concerned with the methods and effectiveness of teaching) indicate that lectures by themselves are a poor way of engaging students and promoting learning. To address this problem, this course will use a Blended Learning/Hybrid Classroom Format. This involves shifting the majority of the material presented in class to out of class. Instructional core content is delivered online, outside of the classroom. Class time is spent exploring topics in greater depth and creates meaningful learning opportunities. This rearrangement allows for more interactive, active learning opportunities during class time like group discussion, Q&A, problem solving activities, and labs where students will apply the methods presented to real data. It also allows for self-paced comprehension of highly complex core concepts.

*Biomedical big data seminar: *Fall 2019, *Fall 2020, *Spring 2021.

*Targeted Learning in Practice PH 243B: *Spring 2021, *Spring 2022.

Consulting in Causal Inference Seminar: Fall and Spring 2009, 2010, 2011, 2012, 2013

Seminar in Genomics plus Lectures on Cross-validation methodology; Spring 2003

Computational Biology Seminar; Spring 2000

DISSERTATION ADVISING

1. Alan Hubbard, UC Berkeley (Biostatistics, 1998), “Applications of Locally Efficient Estimation in Censored Data Models”. Biostatistics Student of the year, 1998, and received the Evelyn Fix Award from the Department of Statistics. Alan Hubbard is a Professor at the Division of Biostatistics, University of California, Berkeley.
2. Derick Peterson (Biostatistics, 1998), “On Nonparametric Estimation and Inference with Censored Data, Bandwidth Selection for Local Polynomial Regression, and Subset Selection in Explanatory Regression”. Student of the year, 1998. Derick Peterson is Associate Professor, Department of Biostatistics, School of Medicine, University of Rochester, Rochester.
3. Chris Quale (Biostatistics, 2001), “Estimation of the Bivariate Survival Function with Censored Truncated Data and Hazard Estimation Based on Interval Censored data”. Senior Quantitative Analyst at Google.
4. Maja Pavlic (Biostatistics, 2001), “Statistical Methods for Analysis of Recurrent Event Data” and “Estimation of the Number of Components in a Mixture of Normals”. Funded by Biotech Company Genentech. Student of the year, 2001. Director at Pharmaceutical Company Johnson Johnson.
5. Jennifer Bryan (Biostatistics, 2001), “Statistical Inference for Gene-expression Analysis from cDNA Microarrays”, Biostatistics Student of the year, 2001, and received the Evelyn Fix Award from the Department of Statistics. Jennifer Bryan is Professor in the Statistics Department and the Department of Biotechnology at the University of British Columbia, Vancouver.
6. Tanya Henneman (Biostatistics, 2002), “Causal Inference in Point Treatment Studies with Applications” funded by Chancellors Opportunity Scholarship, Student of the Year, 2002. Tanya Henneman works in public health department, Berkeley, CA.

7. Alan Brookhart (Biostatistics, 2003), “Computer Intensive Methods in Statistics”, Alan Brookhart was Instructor of Medicine, Harvard Medical School, and Biostatistician, Division of Pharmacoepidemiology and Pharmacoeconomics Brigham and Women’s Hospital. Currently, he is Professor in Biostatistics and Epidemiology at UNC, Chapel Hill.
8. Sunduz Keles (Biostatistics, 2003), “Statistical Methods for Detection of cis-regulatory binding sites”, “Double robust estimation of the Bivariate Survival Function in longitudinal studies”, “Model selection in regression for censored data”. Student of the year, 2003, Public Health Award 2003. Sunduz Keles is Professor, Department of Biostatistics and Statistics, University of Wisconsin, Madison.
9. Katherine Pollard (Biostatistics, 2003), “Computationally Intensive Statistical Methods for Analysis of Gene Expression Data”. Student of the year, 2003, and received the Evelyn Fix Award from the Department of Statistics. Katherine Pollard is Professor in genomics at the Gladstone Institute UCSF.
10. Zhuo Yu (Statistics, 2003), “Causal inference in longitudinal studies”. Received the Erich Lehmann Award from the Department of Statistics, 2003. Zhuo Yu is a Quantitative Analyst at Morgan Stanley, and former research biostatistician at Bristol-Myers Squibb company.
11. Annette Molinaro, (Biostatistics, 2004). “Data Adaptive Prediction in Cancer Research”. Biostatistics Student of the year 2004 and received the Evelyn Fix Award from the Department of Statistics. Funded by grant from the Lawrence Livermore National Laboratory which provided access to their super-computers. Associate Professor in Residence of Neurological Surgery and Epidemiology and Biostatistics at UCSF.
12. Romain Neugebauer, (Biostatistics, 2004). “Double Robust Estimation in Causal Inference Models and its Application in the Analysis of Longitudinal Air Pollution Studies”. Biostatistics Student of the year 2004, and received the Erich Lehmann Award from the Department of Statistics. Funded by Prof. I. Tager (epidemiology grant) and LSI/Chiron grant. Senior Statistician at Kaiser Permanente.
13. Biao Xing, (Biostatistics, 2005). “Fitting Multinomial mixtures to Detect cis-Regulatory Binding Sites and Pathway Analysis in Computational Biology”. Associate Director of Biostatistics at Onyx (Amgen) Pharmaceuticals.
14. Yue Wang, (Biostatistics, 2006). “Data Adaptive Estimation in Causal Inference”. Bristol Meyers.
15. Sandra Sinisi, “Data Adaptive Prediction with the Deletion/Substitution/Addition Algorithm: Applications in Genomics”. Sr. Principal Statistician at Novartis.
16. Merrill Birkner, (Biostatistics, 2006). “Statistical methods for Genomic data”. Funded by Genomics Training grant. Associate Director, Portfolio Management & Operations at Genentech.
17. Maya Petersen (joint with Prof. Art Reingold, Epidemiology, 2006), Causal effects of dynamic treatment interventions with applications in HIV research. Associate Professor Department of Biostatistics and Epidemiology at UC Berkeley.
18. Ed Bein, (Biostatistics, 2006, joint with Alan Hubbard). graduated Fall 2006. Independent Education Management Professional.

19. Oliver Bembom, (Biostatistics, 2008). “Causal inference for realistic rules”. Senior Data scientist at variety of companies, currently at Pandora (2014).
20. Dan Rubin, (Biostatistics, 2009). Double Robust Estimation. Senior Statistician FDA, Safety analysis group.
21. Kelly Moore, (Biostatistics, 2009). “Targeted Maximum Likelihood in Clinical Trials and Safety Analysis”. Senior statistician at Gap Inc./Growth, Innovation, Digital (GID), a division of Gap, Inc..
22. Cathy Tuglus, (Biostatistics, 2010). “Targeted Maximum Likelihood and Variable Importance Analysis”. Biostatistics manager at Amgen.
23. Eric Polley, (Biostatistics, 2010). “Super Learning”. Senior Statistician NIH cancer center, Biostatistician Mayo Clinic.
24. Ori Stittelman, (Biostatistics, 2010). “Collaborative Targeted MLE of causal effect of treatment on time until event outcomes”. Data scientist at variety of companies (e.g Wells Fargo, Media6Degrees), Senior Data Scientist at Dstillery (2014).
25. Sherri Rose, (Biostatistics, 2011). “Targeted Maximum Likelihood and Case Control Data”. Assistant Professor Harvard School of Public Health.
26. Susan Gruber, (Biostatistics, 2011). “Collaborative Targeted MLE, and clinical trials”. Visiting Scientist at Department of Epidemiology, Harvard, and Senior Director for Methods Research at Innovation in Medical Evidence Development and Surveillance (FDA).
27. Kristin Porter, (Biostatistics, 2011). Genomics Training Grant. Senior statistician at MDRC.
28. Jordan Brooks, (Biostatistics, 2012). Received Erich Lehmann Award. Medical researcher, biostatistics, Strauss & Shavelle, Inc.
29. Ivan Diaz, (Biostatistics, 2013). “Causal effects of continuous exposures”. Post-doc at Department of Biostatistics at Johns Hopkins University. Statistician Google, Assistant Professor Biostatistics Cornell.
30. Paul Chaffee, (Biostatistics, 2013). “Targeted MLE in sequentially randomized controlled trials, dynamic treatments”. Sr. Statistician at JustAnswer.com.
31. Wenjing Zheng, (Biostatistics, 2014). “Adaptive Designs, Direct effect, Asymptotics of Cross-validated Targeted MLE”. CAPS Postdoc UCSF, network analysis. Senior Statistician at Netflix.
32. Sam Lendle, (Biostatistics, 2015). “Targeted Learning in Safety Analysis”. Funded by Kaiser Permanente, Senior Statistician at Pandora.
33. Stephanie Sapp, (Biostatistics, 2014). “Big Data”, graduated May, 2014. Senior Statistician at Google.
34. Boriska Toth, (Biostatistics, 2016). “Instrumental Variables”. Kaiser Permanente Oakland.
35. Dan Brown, (Biostatistics, 2014). “Causal effects of exposure to environmental agents during jobs”. SPH UC Berkeley staff member.
36. Laura Balzer, (Biostatistics, 2015). SEARCH trial. Laura is a postdoc in Biostatistics, and the Harvard School of Public Health. Assistant Professor Biostatistics, UmassAmherst.

37. Molly Davies, (Biostatistics, 2015). “Statistical inference for dependent data”. The Climate Corporation, and started data scientist Stitch Fix (2018).
38. Oleg Sofrygin, (Biostatistics, 2016). ”Causal inference for networks”, statistician at Kaiser Permanente. Postdoc on grant from Gates Foundation, UC Berkeley. Kaiser Permanente (2018).
39. Alex Luedtke, (Biostatistics, 2016). Alex is an Assistant Professor at the Fred Hutchinson Cancer Research Center, and now Statistics Department (2018).
40. Erin LeDell, (Biostatistics, 2015). Erin works at H2O as professional software developer in machine learning, received Lehmann award.
41. Cheng Ju, started Fall 2015, graduated June 2018. Collaborative learning for continuous tuning. Started at Facebook.
42. *Aurelien Florent Bibaut, started Fall 2015, Graduated 2020. Senior statistician Netflix.
43. *Mary Combs, starts Fall 2016.
44. * Wilson Cai, started Fall 2016, graduated Fall 2019, quantitative researcher, Citadel.
45. *Ivana Malanica, started Fall 2016, graduated Fall 2022.
46. *Jonathan Levy, started Fall 2015, graduated Fall 2019, Genentech senior statistician.
47. *Chi Zhang, started Fall 2015, graduated Fall 2019, works at data science start up company.
48. * Nima Jehazi, started Fall 2019, graduated summer 2022. Assistant Professor, Department of Biostatistics, Harvard University.
49. *Rachael Phillips, started Fall 2019.
50. *Yue You, started Fall 2019, graduated Spring 2022.
51. * George Shan, started Fall 2019.
52. *William Krinsman, Fall 2018-2022. Graduated Spring 2022.
53. Advised Ph.D student Helene Rytgaard, University of Copenhagen, who visited Spring 2018.
54. Advised Ph.D student Frank Blauw, University of Groningen, who visited Spring 2017.
55. * Advised Ph.D student Oliver Hines visiting for semester from Biostatistics, London School of Hygiene.
56. *Lauren Dang, will graduate Fall 2022.
57. * Haodong Li, Fall 2021-
58. * Yi Li, Fall 2021-
59. * David McCoy, Fall 2020-
60. *Wenxin Zhang, Fall 2021-
61. *Kirsten Landsiedel, Spring 2022-
62. *David Chen, Fall 2020-

POST DOCTORAL ADVISING

1. Chris Andrews (1997-1998), NSF Postdoctoral Fellow in Biostatistics.
2. Joerg Rahnenfuhrer (2001-2002, Germany), Postdoc in Biostatistics.
3. Chris Andrews (2001-2002), NSF Postdoctoral Fellow in Biostatistics.
4. Jonas Larson (2002-2003), Denmark Postdoctoral Fellowship.
5. Sunduz Keles (2003-2004), Postdoctoral Fellow NIH Genomics Grant (joint with Sandrine Dudoit).
6. Blythe Durbin (2003-2005), Postdoctoral Fellow (joint with Sandrine Dudoit).
7. Supervise (jointly with Ira Tager) Romain Neugebauer (2004-2007), Assistant Researcher.
8. Michael Rosenblum (2006-2008), Postdoctoral Fellow. Associate Professor Johns Hopkins.
9. Advisor of visitor Ph.D student G. Reeves (2007), Free University of Amsterdam.
10. Hui Wang (2007-2009), Postdoctoral Fellow.
11. Professor Mathias Drton (co-advisor with Sandrine Dudoit, Lior Pachter, and Bernd Sturmfels), Department of Mathematics, UC Berkeley (Summer 2004-Spring 2005). Project: Multiple testing procedures in graphical model selection. Current position: Assistant Professor, Department of Statistics, University of Chicago.
12. 2011-2013: Marco Carone, Post-doctoral fellow. Assistant Professor University of Washington.
13. 2012-2013, Romain Pirrachio, Post-doctoral fellow.
14. *Mentor of Post-doctoral clinic fellow at UCSF, Carina Marquez, 2014-
15. *Mentor of Gilmer Valdez K-Award (2020-).
16. Post-doctoral fellow, Wenjing Zheng, (2014-2017)
17. *Mentor of Ben Arnold on K-award (2014-)
18. Post-doctoral fellow, Kara Rudolph, (2014-2017).
19. Mentor of clinical fellow Simon Pollett (2014-)
20. Post-doctoral fellow Caleb Miles (2015-2018). Assistant Professor Columbia University.
21. Post-doctoral fellow David Benkezer (2016-2018). Assistant Professor Emery University.
22. Post-doctoral fellow Oleg Sofrygin (2016-2018).
23. Mentor/Advised Assistant-Professor Karla Diaz-Ordaz, London Empirical College, who visited Fall 2016-Spring 2017 to learn about targeted learning.
24. Mentor/Advised Assistant-Professor Jenny Hagstrom, Department of Statistics, USBE, Ume University, Sweden, who visited Fall 2016, to learn about targeted learning.
25. Postdoctoral fellow Benoit Lepage from Paris, (2016-2017).
26. *Postdoctoral fellow, Melih Agraz, 2019.
27. *Postdoctoral fellow, Zeyi Wang, 2020.
28. *Postdoctoral fellow, Aaron Hudson, 2020-2022.

29. *Postdoctoral fellow Alejandro Schuler, 2021-2022.

BIBLIOGRAPHY

PUBLISHED ARTICLES

1. M.J. van der Laan (1994), Modified EM-estimator of the Bivariate Survival Function. *Mathematical Methods of Statistics* 3, 213–43.
2. M.J. van der Laan (1995), An Identity for the Nonparametric Maximum Likelihood Estimator in Missing Data and Biased Sampling Models. *Bernoulli* 1, 4, pp. 335–41.
3. R.D. Gill, M.J. van der Laan, J.A. Wellner (1995), Inefficient Estimators of the Bivariate Survival Function for Three Models. *Annales de L'I.H.P. Prob. Stat.* 31, 3, 545–97.
4. N.P. Jewell, M.J. van der Laan (1995), Generalizations of Current Status Data with Applications. *Lifetime Data Analysis* 1, 101–109.
5. M.J. van der Laan (1996), Efficient Estimation in the Bivariate Censoring Model and Repairing NPMLE. *Annals of Statistics* 24, 2, 596–627.
6. M.J. van der Laan (1996), Nonparametric Estimation of the Bivariate Survival Function with Truncated Data. *Journal of Multivariate Analysis* 58, 1, 107–131.
7. M.J. van der Laan (1996), Efficiency of the NPMLE in the Line-Segment Problem. *Scand. J. Statist.* 23, 527–50.
8. M.J. van der Laan (1996), Efficient and ad hoc Estimation in the Bivariate Censoring Model. *Proceedings of the 1994 Conference on Lifetime Data Models in Reliability and Survival Analysis*, 339–346. Refereed.
9. M.J. van der Laan (1997), Nonparametric Estimators of the Bivariate Survival Function under Random Censoring. *Statistica Neerlandica* 51, 2, 178–200.
10. R.D. Gill, M.J. van der Laan, J.R. Robins (1997), Coarsening at Random: Characterizations, Conjectures and Counter-Examples. *Proceedings of the First Seattle Symposium in Biostatistics*, 1995. D.Y. Lin and T.R. Fleming (editors), Springer Lecture Notes in Statistics, 255–294 (Refereed).
11. N.P. Jewell, M.J. van der Laan (1997), Singly and Doubly Censored Current Status Data with Extensions to Multi-State Counting Processes. *Proceedings of the First Seattle Symposium in Biostatistics*, 1995. D.Y. Lin and T.R. Fleming (editors), Springer Lecture Notes in Statistics, 171–184 (Refereed).
12. M.J. van der Laan (1997), Book review of PROBABILISTIC CAUSALITY IN LONGITUDINAL STUDIES by Mervi Eerola, Springer-Verlag, New York, 1994, *Statistics in Medicine* 16, 23, 2761–62.
13. M.J. van der Laan, P.J. Bickel, N.P. Jewell (1997), Singly and Doubly Censored Current Status Data: Estimation, Asymptotics, Regression. *Scandinavian Journal of Statistics* 24, 289–307.
14. M.J. van der Laan, N.P. Jewell, D. Peterson (1997), Efficient Estimation of the Lifetime and Disease Onset Distribution. *Biometrika* 84, 3, 539–554.

15. M.J. van der Laan, I. McKeague (1997), Efficient Estimation from Right-Censored Data when Failure Indicators are Missing at Random. *Annals of Statistics* 26 164–82.
16. M.J. van der Laan, A. Hubbard (1997), Estimation with Interval Censored Data and Covariates. *Lifetime Data Models* 3, 77–91.
17. M.J. van der Laan (1998), Identity for NPMLE in Censored Data Models, *Lifetime Data Models* 4, 83–102.
18. M.J. van der Laan (1998), The Two-Interval Line-Segment Problem. *Scandinavian Journal of Statistics* 25, 163–86.
19. M.J. van der Laan, A. Hubbard (1998), Locally Efficient Estimation of the Survival Distribution with Right Censored Data and Covariates when Collection of Data is Delayed. *Biometrika* 85, 4, pp. 771–83.
20. M.J. van der Laan, J.M. Robins (1998), Locally Efficient Estimation with Current Status Data and Time-Dependent Covariates. *Journal of the American Statistical Association* 93, 442, 693–701.
21. A. Hubbard, M.J. van der Laan, J.M. Robins (1999), Nonparametric locally efficient estimation of the treatment specific survival distribution with right censored data and covariates in observational studies, *Statistical Models in Epidemiology, The Environment and Clinical trials*, IMA Volumes in Mathematics and its Applications, Ed. M.E. Halloran and D. Berry, Springer Verlag, Vol. 116, 135–178.
22. M.J. van der Laan, A. Hubbard (1999), Locally efficient estimation of the quality adjusted lifetime distribution with right-censored data and covariates, *Biometrics* 55, 530–36.
22. M.J. van der Laan, R.D. Gill (1999), Efficiency of the NPMLE in Nonparametric Missing Data Models. *Mathematical Methods of Statistics* 8, 2, 251–76.
23. M.J. van der Laan (1999), Discussion of ‘Adjusting for Non-ignorable Drop Out Using Semiparametric Non-response Models’ by Scharfstein, Rotnitzky and Robins, *the Journal of the American Statistical Association* 94, 448, 1125–1128.
24. J.M. Robins, A. Rotnitzky and M.J. van der Laan (1999), Discussion of ‘On Profile Likelihood’ by S.A. Murphy and A.W. van der Vaart, *Journal of the American Statistical Association* 95, 477–82.
25. M.J. van der Laan, C. Andrews (2000), The Nonparametric Maximum Likelihood Estimator in a class of doubly censored current status data models with application to partner studies, *Biometrika* 87, 61–71. 1410-1424
26. M.J. van der Laan, P. van der Laan (2000), Subset selection based on order statistics from logistic populations, *Statistics* 00, 1–9.
27. A.E. Hubbard, M.J. van der Laan, W. Enanoria, J. Colford (2000), Nonparametric Survival Estimation When Death is Reported with Delay, *Lifetime Data Models* 6, 237–50
28. C. Quale, M.J. van der Laan (2000), Inference with Bivariate Truncated Data, *Lifetime Data Analysis* 6, 4, 391–408.
29. C. Quale, M.J. van der Laan, J.M. Robins (2006), Locally efficient estimation with bivariate right censored data, *Journal of the American Statistical Association*
30. M.J. van der Laan, J. Bryan (2001), Gene Expression Analysis with the Parametric Bootstrap, *Biostatistics* 2, 3, 1–17.

31. M.J. van der Laan, N.P. Jewell (2001), The NPMLE in the Uniform Doubly Censored Current Status Data Model, *Scandinavian Journal of Statistics* 28, 537–549.
32. S.A. Murphy, M.J. van der Laan, J.M. Robins (2001), Marginal Mean Models for Dynamic Treatment Regimes, *Journal of the American Statistical Association* 96, 1410–1424.
33. M.J. van der Laan, Zhuo, Y. (2001), Comments on the millenium paper 'Inference for semiparametric models: Some questions and an answer', by P.J. Bickel and J. Kwon, in the *millennium series of Statistica Sinica*, 910–917.
34. K. Pollard, M.J. van der Laan (2002), Statistical Inference for Simultaneous Clustering of Gene Expression Data, *Journal of Mathematical Biosciences* 176, 1, 99–121.
35. J. Bryan, K. Pollard, M.J. van der Laan (2002), Paired and Unpaired Comparison and Clustering with Gene Expression Data, Special issue on Bioinformatics in *Statistica Sinica* 12, 1, 87–110.
36. M.J. van der Laan, A. Hubbard, J.M. Robins (2002), Locally Efficient Estimation of a Multivariate Survival Function in Longitudinal Studies, *Journal of the American Statistical Association* 97, 494–508.
37. C. Johnstone, T. Henneman, C. McCulloch, M.J. van der Laan (2002), Modeling Treatment Effects on Binary Outcomes with Grouped-Treatment Variables and Individual Covariates, *American Journal of Epidemiology* 156, 753–60.
38. M.J. van der Laan, A.W. van der Vaart (2002), Smooth Estimation of a monotone density *Statistics* 37, 3, 189–203.
39. S. Keles, M.J. van der Laan, M. Eisen (2002), Identification of Regulatory Elements Using A Feature Selection Method, *Bioinformatics* 18, 1167–1175.
40. M. A. Brookhart, A. E. Hubbard, M. J. van der Laan, J. M. Colford, J.N.S. Eisenberg (2002). Statistical Estimation of Parameters in a Disease Transmission Model: Analysis of a *Cryptosporidium* Outbreak. *Statistics in Medicine* 21, 23, 3627–3638.
41. Z. Yu, M.J. van der Laan (2005), Construction of counterfactuals and the G-computation formula, *Mathematical Methods of Statistics*, vol. 14, no. 4 (2005), 488–499.
42. M.J. van der Laan, N.P. Jewell (2003), Current Status and Right-Censored Data Structures when Observing a Marker at the Censoring Time, *Annals of Statistics* 31, 2, 512–35.
43. M. Miloslavsky, M.J. van der Laan (2003), Fitting of Mixtures with Unspecified Number of Components using Cross-Validation Distance Estimate, *Computational Statistics and Data Analysis* 41, 413–428.
44. N.P. Jewell, M.J. van der Laan, T. Henneman (2003), Nonparametric Estimation from Current Status Data with Competing Risks, *Biometrika* 90, 1, 183–97.
45. Jean C Norris, Mark J Van der laan, Sylvia Lane, James N Anderson, and Gladys Block, Nonlinearity in Demographic and Behavioral Determinants of Morbidity Health Serv Res. 2003 December; 38(6 Pt 2): 1791-1818. doi: 10.1111/j.1475-6773.2003.00203.
46. M.J. van der Laan, K.S. Pollard (2003), A New Algorithm for Hierarchical Hybrid Clustering with Visualization and the Bootstrap, *Journal of Statistical Planning and Inference* 117, 275–303.

47. M.J. van der Laan, K. Pollard, J. Bryan (2003), A new partitioning around medoids algorithm, *Journal of Statistical Computation and Simulation* 73, No. 8, 575–584.
48. M.J. van der Laan, A.W. van der Vaart (2006), Estimating a Survival Distribution with Current Status Data and High-Dimensional Covariates, *International Journal of Biostatistics*, Vol. 2 : Iss. 1, Article 9. Available at: <http://www.bepress.com/ijb/vol2/iss1/9>
49. M. Miloslavsky, S. Keles, M.J. van der Laan, S. Butler (2003), Recurrent event analysis in the presence of time-dependent covariates and dependent censoring, *Journal of the Royal Statistical Society, Series B*, 66, Part 1, 239–257.
50. S. Keles, M.J. van der Laan, S. Dudoit, B. Xing, M. B. Eisen (2003), Supervised detection of regulatory motifs in DNA sequences, *Statistical Applications in Genetics and Molecular Biology* 2, 1, Article 5. <http://www.bepress.com/sagmb/vol2/iss1/art5S>.
51. S. Dudoit, M. J. van der Laan, S. Keles, A. M. Molinaro, S. E. Sinisi, S. L. Teng (2003). Loss-based estimation with cross-validation: Applications to microarray data analysis. In G. Piatetsky-Shapiro and P. Tamayo (eds), *Microarray Data Mining*, Special Issue of SIGKDD Explorations, Vol. 5, No. 2, p. 56-68.
52. S. Keles, M.J. van der Laan, J.M. Robins (2004), Estimation of the bivariate survival function with generalized bivariate right censored data structures, Chapter 8, pages 143-175, in *Advances in Survival Analysis*, Edited by N. Balakrishnan and C.R. Rao, *Handbook of Statistics 23*, Elsevier North Holland.
53. N.P. Jewell, M.J. van der Laan (2004), Current status data: review, recent developments and open problems, Chapter 35, pages 625-643, in *Advances in Survival Analysis*, Edited by N. Balakrishnan and C.R. Rao, *Handbook of Statistics 23*, Elsevier North Holland.
54. K. Pollard, M.J. van der Laan (2004), Choice of null distribution in resampling based multiple testing, *Journal of Statistical Planning and Inference* 125, 85–101.
55. A. M. Molinaro, S. Dudoit, M. J. van der Laan (2004). Tree-based multivariate regression and density estimation with right-Censored data. In S. Dudoit, R. C. Gentleman, and M. J. van der Laan (eds), Special Issue on Multivariate Methods in Genomic Data Analysis, *Journal of Multivariate Analysis* 90, 1, p. 154–77.
56. N.P. Jewell, M.J. van der Laan (2004), Case control current status data, *Biometrika*, **91**, 3, 529-541.
57. J. Bryan, Z. Yu, M.J. van der Laan (2004), Analysis of longitudinal marginal structural models, *Biostatistics* 5, 3, pp. 361–80
58. M. J. van der Laan, S. Dudoit, S. Keles (2004), Asymptotic optimality of likelihood-based cross-validation, *Statistical Applications in Genetics and Molecular Biology* 3, 1, Article 4. <http://www.bepress.com/sagmb/vol3/iss1/art4>.
59. M.J. van der Laan, S. Dudoit, K.S. Pollard (2004), Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives, *Statistical Applications in Genetics and Molecular Biology* 3, 1, Article 15. <http://www.bepress.com/sagmb/vol3/iss1/art15>
60. S. Keles, M. J. van der Laan, and S. Dudoit (2004), Asymptotically Optimal Model Selection Method for Regression on Censored Outcomes, *Bernoulli* 10, 6, 1011-1037.

61. S. Dudoit, M.J. van der Laan, K.S Pollard (2004), Multiple testing. Part I. Single-step procedures for control of general Type I error rates, *Statistical Applications in Genetics and Molecular Biology* Vol. 3: No. 1, Article 13. <http://www.bepress.com/sagmb/vol3/iss1/art13>
62. M.J. van der Laan, S. Dudoit, K.S Pollard (2004), Multiple testing. Part II. Step-down procedures for control of the family-wise error rate, *Statistical Applications in Genetics and Molecular Biology* Vol. 3: No. 1, Article 14. <http://www.bepress.com/sagmb/vol3/iss1/art14>
63. I.B Tager, T. Haight, B. Sternfeld, Z. Zhou, M.J. van der Laan (2004), Effects of physical activity and body composition on functional limitation in the elderly: Application of the marginal structural model, *Epidemiology* 15, 479–93.
64. S. Keles, M.J. van der Laan, C. Vulpe (2004), Regulatory Motif Finding by Logic Regression, *Bioinformatics* 20, 2799–2811.
65. S. Sinisi, M.J. van der Laan (2004), The deletion/substitution/addition algorithm in loss function based estimation: Applications in Genomics, *Journal of Statistical Methods in Molecular Biology*, Vol. 3, No. 1, Article 18, <http://www.bepress.com/sagmb/vol3/iss1/art18>.
66. R. Neugebauer, M.J. van der Laan (2005) Why prefer double robust estimators in causal inference? *Journal of Statistical Planning and Inference*, Volume 129, Issues 1-2, 15 February 2005, Pages 405-426.
67. B. Xing, M.J. van der Laan (2005), A statistical method for constructing transcriptional regulatory networks using gene expression and sequence data, *Journal of Computational Biology* 12, 2, 229–246.
68. C. Andrews, M.J. van der Laan, J.M. Robins (2005), Locally Efficient Estimation of Regression Parameters Using Current Status Data, *Journal of Multivariate Analysis* 96, 2, 332–51.
69. S. Dudoit, M.J. van der Laan (2005), Asymptotics of cross-validated risk estimation in estimator selection and performance assessment. *Statistical Methodology* 2, 2, 131–54.
70. K. Mortimer, R. Neugebauer, M.J. van der Laan, I.B. Tager (2005), An application of model fitting procedures for marginal structural Models, *American Journal of Epidemiology* 162, 607–17.
71. T. Haight, I. Tager, B. Sternfeld, W. Satariano, M. van der Laan (2005), Effects of body composition and leisure-time physical activity on transitions in physical functioning in the elderly. *The American Journal of Epidemiology* 162, 607–17.
72. M. van der Laan, T. Haight, I. Tager (2005), Discussion: Hypothetical interventions to define causal effects: afterthought or prerequisite? *The American Journal of Epidemiology* 162, 382–88.
73. Nicholas P. Jewell, M. van der Laan and X. Lei, Bivariate current status data with univariate monitoring times, *Biometrika*, 92, 2005, 847-862.
74. M. D. Birkner, S. E. Sinisi, M. J. van der Laan (2005), Multiple Testing and Data Adaptive Regression: An Application to HIV-1 Sequence Data, *Statistical Applications in Genetics and Molecular Biology* 4, 1, Article 8.

75. K.S. Pollard, S. Dudoit, M.J. van der Laan, MJ (2005), Multiple testing procedures: the multtest package and applications to genomics. Chapter 15 (pages 249–271) in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer, New York.
76. K.S. Pollard, M.J. van der Laan (2005), Cluster analysis of genomic data with applications in R, in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer.
77. B. Xing, M.J. van der Laan (2005), A causal inference approach for constructing transcriptional regulatory networks, *Bioinformatics* 21, 4007–13.
78. M J. van der Laan, M.D. Birkner, A.E. Hubbard (2005), Empirical Bayes and resampling based multiple testing procedure controlling tail probability of the proportion of false positives, *Statistical Applications in Genetics and Molecular Biology* 4, 1, Article 29.
79. R. Neugebauer, M. J. van der Laan (2007), Nonparametric Causal Effects based on marginal structural models, *Journal of Statistical Planning and Inference* 137, 419–434.
80. M. Petersen, S. E. Sinisi, M.J. van der Laan (2006), Estimation of direct causal effects, *Epidemiology* 17(3): 276-284.
81. Mark J. van der Laan, Maya L. Petersen, and Marshall M. Joffe (2005) "History-Adjusted Marginal Structural Models and Statically-Optimal Dynamic Treatment Regimens ", The International Journal of Biostatistics: Vol. 1: No. 1, Article 4. <http://www.bepress.com/ijb/vol1/iss1/4>.
82. T. Hothorn, P. Buhlmann, S. Dudoit, A. Molinaro, and M.J. van der Laan (2005), Survival Ensembles, *Biostatistics*, 7, No. 3, 355–373.
83. A. Barrier, A. Lemoine, P.-Y. Boelle, C. Tse, D. Brault, F. Chiappini, J. Breittschneider, F. Lacaine, S. Houry, M. Huguier, M. J. van der Laan, T. P. Speed, B. Debuire, A. Flahault, and S. Dudoit (2005). Colon cancer prognosis prediction by gene expression profiling. *Oncogene*, Vol. 24, No. 40, p. 6155-6164. <http://www.nature.com/onc/journal/v24/n40/index.html>.
84. M.A. Brookhart, M.J. van der Laan (2006), A semiparametric model selection criterion with applications to the marginal structural model, *Journal of Computational Statistics and Data Analysis* 50, 2, 457–98.
85. Z. Yu, M.J. van der Laan (2006), Double Robust Estimation in Longitudinal Marginal Structural Models, *Journal of Statistical Planning and Inference* 136, 3, 1061–89.
86. S. Mukherjee, S.J. Roberst, M.J. van der Laan (2005), Data-adaptive test statistics for microarray data, *Bioinformatics* 21, Supplement 2, ii108-ii114, <http://bioinformatics.oxfordjournals.org/cgi/content/abstract/21/suppl2/ii108>
87. S. Keles, M.J. van der Laan, S. Dudoit, S. Cowley, S.L. Teng (2006), Multiple testing methods for CHIP-Chip high density oligonucleotide array data. *Journal of Computational Biology*, Vol. 13, No. 3, p. 579-613. [<http://www.liebertonline.com/doi/abs/10.1089/cmb.2006.13.579>] [Tech report #147].
88. Mark J. van der Laan (2006) "Statistical Inference for Variable Importance", The International Journal of Biostatistics: Vol. 2: No. 1, Article 2. <http://www.bepress.com/ijb/vol2/iss1/2>

89. Mark J. van der Laan and Daniel Rubin (2006), Estimating Function Based Cross-Validation and Learning, U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 180. <http://www.bepress.com/ucbbiostat/paper180>, to appear in *Frontiers of Statistics*, Springer, Festschrift in honour of 65-th Birthday of Peter Bickel.
90. Mark J. van der Laan and Alan E. Hubbard (2006) "Quantile-Function Based Null Distribution in Resampling Based Multiple Testing," *Statistical Applications in Genetics and Molecular Biology*: Vol. 5: No. 1, Article 14. Available at: <http://www.bepress.com/sagmb/vol5/iss1/art14>.
91. Merrill D. Birkner, Alan E. Hubbard, Mark J. van der Laan, Christine F. Skibola, Christine M. Hegedus, and Martyn T. Smith (2006), "Issues of Processing and Multiple Testing of SELDI-TOF MS Proteomic Data", *Statistical Applications in Genetics and Molecular Biology*, Vol. 5, Issue 1: <http://www.bepress.com/sagmb/vol5/iss1/art11>. Also U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 200. <http://www.bepress.com/ucbbiostat/paper200>
92. Y. Wang, O. Bembom, M.J. van der Laan (2007), Data-Adaptive Estimation of the Treatment-Specific Mean. *Journal of Statistical Planning and Inference*, 137(6):1871-1887.
93. V. De Gruttolas, C. Flexner, D. Kuritzkes, J. Leland, M. Hughes (2006), Drug Development Strategies for Salvage Therapy: Conflicts and Solutions, *AIDS Research and Human Retroviruses*.
94. B. Durbin, S. Dudoit, M.J. van der Laan (2006), A Deletion/Substitution/Addition Algorithm for Classification Neural Networks, with Applications to Biomedical Data", *Journal of Statistical Planning and Inference* **138**, No 2., 464-488.
95. Daniel Rubin, Sandrine Dudoit, and Mark van der Laan (2006) "A Method to Increase the Power of Multiple Testing Procedures Through Sample Splitting," *Statistical Applications in Genetics and Molecular Biology*: Vol. 5: No. 1, Article 19. Available at: <http://www.bepress.com/sagmb/vol5/iss1/art19>
96. Sandra E. Sinisi, Romain Neugebauer, and Mark J. van der Laan (2006) "Cross-Validated Bagged Prediction of Survival," *Statistical Applications in Genetics and Molecular Biology*: Vol. 5: No. 1, Article 12. Available at: <http://www.bepress.com/sagmb/vol5/iss1/art12>
97. Merrill D. Birkner and Mark J. van der Laan (2006) "Application of a Variable Importance Measure Method," *The International Journal of Biostatistics*: Vol. 2: No. 1, Article 6. Available at: <http://www.bepress.com/ijb/vol2/iss1/6>
98. Mark J. van der Laan (2006) "Statistical Inference for Variable Importance," *The International Journal of Biostatistics*: Vol. 2: No. 1, Article 2. Available at: <http://www.bepress.com/ijb/vol2/iss1/2>
99. Aad and van der Laan, Mark J. (2006) "Estimating a Survival Distribution with Current Status Data and High-dimensional Covariates," *The International Journal of Biostatistics*: Vol. 2 : Iss. 1, Article 9, Available at: <http://www.bepress.com/ijb/vol2/iss1/9>
100. Nicholas P. Jewell, Mark J. van der Laan, and Stephen Shiboski (2006) "Choice of Monitoring Mechanism for Optimal Nonparametric Functional Estimation for Binary

- Data,” The International Journal of Biostatistics: Vol. 2: No. 1, Article 7. Available at: <http://www.bepress.com/ijb/vol2/iss1/7>
101. Mark J. van der Laan, Sandrine Dudoit, Aad W. van der Vaart (2006), ”The Cross-validated Adaptive Epsilon-Net Estimator”, *Statistics and Decisions*, **24**, No. 3, 373–395.
 102. Aad W. van der Vaart, S. Dudoit, M.J. van der Laan (2006), ”Oracle Inequalities for Multi-Fold Cross-Validation”, *Statistics and Decisions*, **24**, No. 3, 351–371.
 103. Multivariate Methods in Genomic Data Analysis. Special Issue of the Journal of Multivariate Analysis. Edited by S. Dudoit, R.C. Gentleman, and M.J. van der Laan.
 104. K. S. Pollard, M. D. Birkner, M. J. van der Laan, and S. Dudoit (2005). Test statistics null distributions in multiple testing: Simulation studies and applications to genomics. Numero double special Statistique et Biopuces, Journal de la Societe Francaise de Statistique, Vol. 146, No. 1-2, p. 77-115. [Tech report #184].
 105. Romain Neugebauer, M.J. van der Laan (2006), G-computation estimation for causal inference with complex longitudinal data, *Computational Statistics & Data Analysis*, **51**, 1676–1697.
 106. Romain Neugebauer and Mark J. van der Laan (2006), Causal effects in longitudinal studies: Definition and maximum likelihood estimation. *Computational Statistics & Data Analysis*, **51**, 1664–1675.
 107. Sandrine Dudoit, Sunduz Keles, Mark van der Laan (2007), Multiple tests of association with biological annotation data base, IMS Lecture Notes-Monograp Series, Probability and Statistics: Essays in Honour of David A. Freedman, Institute of Mathematical Statistics. [Tech report #202], <http://www.bepress.com/ucbbiostat/paper202>.
 108. Maya L. Petersen, Steven G. Deeks, Jeffrey N. Martin, and Mark J. van der Laan, History-Adjusted Marginal Structural Models to Estimate Time-Varying Effect Modification Am J Epi: 2007 166(9):985-93. Discussion paper. Also available as Working Paper Series. Working Paper 199. <http://www.bepress.com/ucbbiostat/paper199>.
 109. Maya Petersen, Mark van der Laan (2007), Response to Commentary by J.M. Robins and Hernan: Effect Modification by Time-Varying Covariates”, American Journal of Epidemiology. History-Adjusted Marginal Structural Models to Estimate Time-Varying Effect Modification Am J Epi: 2007 166(9):985-93.
 110. O. Bombom, M.L. Petersen, M.J. van der Laan (2006), Identifying important explanatory variables for time-varying outcomes. In W. Dubitzky, M. Granzow, and D.P. Berrar (eds.), Fundamentals of Data Mining in Genomics and Proteomics, Springer, Chapter 11, p. 227-250.
 111. M.L. Petersen, Y. Wang, M.J. van der Laan, D. Bangsberg (2006), Assessing the Effectiveness of Antiretroviral Adherence Interventions: Using Marginal Structural Models to Replicate the Findings of Randomized Controlled Trials. JAIDS 2006 43 (Suppl 1): S96-S103.
 112. M.L. Petersen, Steven G. Deeks, Mark J. van der Laan, Individualized Treatment Rules: Generating Candidate Clinical Trials. Stat Med: 2007 26(25):4578-601.
 113. Alan E. Hubbard and Mark J. van der Laan, ”Population Intervention Models in Causal Inference” (October 2005). U.C. Berkeley Division of Biostatistics Working

- Paper Series. Working Paper 191, <http://www.bepress.com/ucbbiostat/paper191>, to appear in *Biometrika*.
114. M.J. van der Laan, A. Hubbard, N.P. Jewell (2007), Estimation of Treatment Effects in Randomized Trials with Noncompliance and a Dichotomous Outcome, technical report, Division of Biostatistics, <http://www.bepress.edu/ucbbiostat>, *Journal of the Royal Statistical Society B (Statistical Methodology)*, **69**: 443–482.
 115. Sandra Sinisi, E. C Polley, Maya Petersen, Soo-Yon Rhee, Mark van der Laan (2007) Super Learning: An Application to the prediction of HIV-1 Drug Resistance, *Statistical Applications in Genetics and Molecular Biology*: Vol. 6 : Iss. 1, Article 7. Available at: <http://www.bepress.com/sagmb/vol6/iss1/art7>
 116. Oliver Bembom, Sunduz Keles, Mark van der Laan (2007), Supervised Detection of Conserved Motifs in DNA Sequences with *cosmo*. *Statistical Applications in Genetics and Molecular Biology*: Vol. 6 : Iss. 1, Article 8. Available at <http://www.bepress.com/sagmb/vol6/iss1/art8>.
 117. Dan Rubin, Mark van der Laan (2007), A Doubly Robust Censoring Unbiased Transformation, *The International Journal of Biostatistics*, Vol. 3 (1), Article 4, available at <http://www.bepress.com/ijb/vol3/iss1/4>.
 118. Mark van der Laan, Maya Petersen (2007), Statistical Learning of Origin-Specific Statically Optimal Individualized Treatment Rules, *The International Journal of Biostatistics*, <http://www.bepress.com/ijb/vol3/iss1/6>
 119. Mark van der Laan, Dan Rubin (2006), Targeted Maximum Likelihood Learning, *The International Journal of Biostatistics*, Vol 2, Iss. 1, Article 11. Available at <http://www.bepress.com/ijb/vol2/iss1/11>.
 120. Maya Petersen, Y. Wang, M.J. van der Laan, J. Fessel, R. Shafer (2007), Virologic Efficacy of Boosted Double vs. Boosted Single Protease Inhibitor Therapy *AIDS*: 2007 21(12): 1547-1554.
 121. W.T.A Enanoria, A.E. Hubbard, M.J. van der Laan, M. Chen, J. Ruiz, J.M. Colford (2007), Early prediction of median survival among a large AIDS surveillance cohort, *BMC Public Health*, 7: 127.
 122. Oliver Bembom, Mark van der Laan (2007), Statistical Methods for analyzing sequentially randomized trials, *Journal of the National Cancer Institute*, 99(21):1577-1582, commentary on JNCI article Adaptive therapy for androgen independent prostate cancer: A Randomized selection trial including four regimens, by Peter F. Thall, C. Logothetis, C. Pagliaro, S. Wen, M.A. Brown, D. Williams, R. Millikan (2007).
 123. M.L. Petersen, A. Molinaro, S.E. Sinisi, M.J. van der Laan (2007) Cross-validated Bagged Learning. *J. Multiv. Analysis*: 2007 98 (9): 1693-1704. Also available as (June 2005), U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 182. <http://www.bepress.com/ucbbiostat/paper182>.
 124. Mark J. van der Laan, Eric Polley, Alan Hubbard (2007), Super Learner, *Statistical Applications in Genetics and Molecular Biology*, <http://www.bepress.com/sagmb/vol6/iss1/art25>.
 125. O. Bembom and M.J. van der Laan (2008). Analyzing sequentially randomized trials based on causal effect models for realistic individualized treatment rules. *UC Berkeley Division of Biostatistics Working Paper Series*. Working Paper 216. Available at: <http://www.bepress.com/ucbbiostat/paper216>, *Statistics in Medicine* 2008 Aug 30; 27(19):3689-716.

126. R. Neugebauer, M. M. Joffe, I. B. Tager and M.J. van der Laan (2007), Causal inference in longitudinal studies with history-restricted marginal structural models. *Electron. J. Statist.* Volume 1, 119-154.
127. J. Chen, M.J. van der Laan, M.T. Smith. A.E. Hubbard (2007), A comparison of methods to control Type I Error in Microarray studies, *Statistical Applications in Genetics and Molecular Biology*: Vol. 6 : Iss. 1, Article 28. Available at: <http://www.bepress.com/sagmb/vol6/iss1/art28>
128. M.L. Petersen, Y. Wang, M.J. van der Laan, D. Guzman, E. Riley, and D.R. Bangsberg (2007), Pillbox Organizers are Associated with Improved Adherence to HIV Antiretroviral Therapy and Viral Suppression: A Marginal Structural Model Analysis. *Clin Infect Dis*: 2007 45(7):908-15.
129. M.J. van der Laan and Maya L. Petersen (2007) "Causal Effect Models for Realistic Individualized Treatment and Intention to Treat Rules," *The International Journal of Biostatistics*: Vol. 3 : Iss. 1, Article 3.
130. Mark J. van der Laan and Maya L. Petersen, Direct Effect Models (2008), *International Journal of Biostatistics*, <http://www.bepress.com/ijb/vol4/iss1/23>
131. D. Rubin and M.J. van der Laan (2007), Empirical Efficiency Maximization: Improved Locally Efficient Covariate Adjustment in Randomized Experiments and Survival Analysis, *The International Journal of Biostatistics*: Vol. 4 : Iss. 1, Article 5.
132. Nancy S Padian, Ariane van der Straten, Gita Ramjee, Tsungai Chipato, Guy de Bruyn, Kelly Blanchard, Stephen Shiboski, Elizabeth T Montgomery, Heidi Fancher, Helen Cheng, Michael Rosenblum, Mark van der Laan, Nicholas Jewell, James McIntyre, the MIRA Team. Diaphragm and lubricant gel for prevention of HIV acquisition in southern African women: a randomised controlled trial. *The Lancet* - Vol. 370, Issue 9583, 21 July 2007, Pages 251-261
133. S. Dudoit, G. Houston, M.J. van der Laan (2008), Resampling-Based Empirical Bayes Multiple Testing Procedures for Controlling Generalized Tail Probability and Expected Value Error Rates: Focus on False Discovery Rate and Simulation Study, *Biometrical Journal*, Vol. 50, issue 5, 716–744.
134. O. Bembom, M.L. Petersen, S.-Y. Rhee , W. J. Fessel, S.E. Sinisi, R.W. Shafer, and M.J. van der Laan (2008). Biomarker discovery using targeted maximum likelihood estimation: Application to the treatment of antiretroviral resistant HIV infection. *UC Berkeley Division of Biostatistics Working Paper Series*. Working Paper 221. Available at: <http://www.bepress.com/ucbbiostat/paper22>. *Statistics in Medicine*: Vol. 28: Iss 1, 152-172.
135. Kelly Moore, M.J. van der Laan (2009), Application of Time-to-Event Methods in the Assessment of Safety in Clinical Trials, Chapter in Book titled *Design, Summarization, Analysis & Interpretation of Clinical Trials with Time-to-Event Endpoints*, Editor Karl E. Peace, Chapman and Hall.
136. Eric Polley, Mark van der Laan (2008), Predicting Optimal Treatment Assignment Based on Prognostic Factors in Cancer Patients, Chapter in book titled *Design, Summarization, Analysis & Interpretation of Clinical Trials with Time-to-Event Endpoints*, Editor Karl E. Peace, Chapman and Hall. .
137. Catherine Tuglus, Mark van der Laan (2008), "Discussion of: Treelets—An adaptive multi-scale basis for sparse unordered data" was published in the *Annals of Applied Statistics* 2008, Vol. 2, No. 2, 489-493.

138. Maya Petersen, Richard Moore, Steve Deeks, Mark van der Laan (2008), 'Long term consequences of the delay between virologic failure of highly active antiretroviral therapy and regimen modification, *AIDS*, 2008, Oct 18;22(16):2097-106.
139. Michael Rosenblum, Mark van der Laan (2008), Using Regression Models to Analyze Randomized Trials: Asymptotically Valid Hypothesis Tests Despite Incorrectly Specified Models. *Biometrics*. 2009; 65(3): 937-945.
140. Kelly Moore and Mark van der Laan (2009), Covariate Adjustment in Randomized Trials with Binary Outcomes: Targeted Maximum Likelihood Estimation, *Statistics in Medicine* 28(1), 39–64.
141. O. Bembom and M.J. van der Laan (2008), Data-adaptive selection of the adjustment set in variable importance estimation, to appear in *Electronic Journal of Applied Statistics*.
142. Mark van der Laan (2008), Estimation Based on Case-Control Designs with Known Prevalence Probability, *The International Journal of Biostatistics: Vol. 4 : Iss. 1*, Article 17.
143. Sherri Rose, Mark van der Laan (2008) "Simple Optimal Weighting of Cases and Controls in Case-Control Studies," *The International Journal of Biostatistics: Vol. 4 : Iss. 1*, Article 19.
144. Michael Rosenblum, Nicholas P. Jewell, Mark J. van der Laan, Stephen Shiboski, Ariane van der Straten, and Nancy Padian (2009), Analyzing Direct Effects in Randomized Trials with Secondary Interventions: An Application to HIV Prevention Trials. *J R Stat Soc Ser A Stat Soc*. 2009 Apr;172(2):443-465. (September 2007). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 223. <http://www.bepress.com/ucbbiostat/paper223>.
145. Michael Rosenblum, Mark J. van der Laan (2009), Confidence Intervals for the Population Mean Tailored to Small Sample Sizes, with Applications to Survey Sampling, *The International Journal of Biostatistics*, 5(1): Article 4.
146. Sherri Rose, Mark J. van der Laan (2009), "Why Match? Investigating Matched Case-Control Study Designs with Causal Effect Estimation," *The International Journal of Biostatistics: Vol. 5 : Iss. 1*, Article 1.
147. Katherine S. Pollard and Mark J. van der Laan (2008), Supervised Distance Matrices, *Statistical Applications in Genetics and Molecular Biology*, <http://www.bepress.com/sagmb/vol7/iss1/art33>.
148. Catherine Tuglus and Mark J. van der Laan (2009), Modified FDR Controlling Procedure for Multi-Stage Analyses SAGMB, 2009 January 1; 8(1): Article 12.
149. O. Bembom and M.J. van der Laan (2007). Estimating the effect of vigorous physical activity on mortality in the elderly based on realistic individualized treatment and intention-to-treat rules, *Electron J Stat*. 2007; 1: 574-596. UC Berkeley Division of Biostatistics Working Paper Series. Working Paper 217. Available at: <http://www.bepress.com/ucbbiostat/paper217>.
150. O. Bembom and M.J. van der Laan (2007). Data-adaptive selection of the truncation level for Inverse-Probability-of-Treatment-Weighted estimators. *Biometrics*.
151. O. Bembom, M.J. van der Laan, T. Haight, and I.B. Tager (2009). Leisure-time physical activity and all-cause mortality in an elderly cohort, *Epidemiology* 20(3): 424-30.

152. Influenza Vaccination and Mortality: Differentiating Vaccine Effects From Bias Bruce Fireman; Janelle Lee; Ned Lewis; Oliver Bembom; Mark van der Laan; Roger Baxter (2009), *Am. J. Epidemiol.* 170 (5): 650-656.
153. Rosenblum M, Deeks SG, van der Laan M, Bangsberg DR (2009) The Risk of Virologic Failure Decreases with Duration of HIV Suppression, at Greater than 50% Adherence to Antiretroviral Therapy. *PLoS ONE* 4(9): e7196. doi:10.1371/journal.pone.0007196
154. Tad Haight, H. Wang, M.J. van der Laan, I. Tager (2010), A cross-validation deletion-substitution-addition model selection algorithm: Application to marginal structural models. *Computational Statistics & Data Analysis* 54(12): 3080-3094 (2010).
155. M.J. van der Laan (2010), Targeted Maximum Likelihood Based Causal Inference Part I, Special Issue on Causal Inference of *International Journal of Biostatistics*, The *International Journal of Biostatistics*: Vol. 6 : Iss. 2, Article 2. DOI: 10.2202/1557-4679.1211 (<http://www.bepress.com/ijb/vol6/iss2/2>).
156. M.J. van der Laan (2010), Targeted Maximum Likelihood Based Causal Inference Part II, Special Issue on Causal Inference of *International Journal of Biostatistics*, The *International Journal of Biostatistics*: Vol. 6 : Iss. 2, Article 3. DOI: 10.2202/1557-4679.1241 (<http://www.bepress.com/ijb/vol6/iss2/3>).
157. M.J. van der Laan, and S. Gruber (2010) "Collaborative Double Robust Targeted Maximum Likelihood Estimation," *The International Journal of Biostatistics*: Vol. 6 : Iss. 1, Article 17.
158. S. Gruber, M.J. van der Laan (2010), An Application of Collaborative Targeted Maximum Likelihood Estimation in Causal Inference and Genomics", *The International Journal of Biostatistics*: Vol. 6 : Iss. 1, Article 18.
159. Rosenblum M, van der Laan MJ. (2010), Targeted Maximum Likelihood Estimation of the Parameter of a Marginal Structural Model, *The International Journal of Biostatistics*: Vol. 6 : Iss. 2, Article 19. DOI: 10.2202/1557-4679.1238. (<http://www.bepress.com/ijb/vol6/iss2/19>).
160. Rosenblum M, van der Laan MJ. (2010), Simple, Efficient Estimators of Treatment Effects in Randomized Trials Using Generalized Linear Models to Leverage Baseline Variables. *International Journal of Biostatistics*. 2010. 6(1): Article 13.
161. A.M. Molinaro, K. Lostritto, M.J. van der Laan (2010), partDSA: Deletion/Substitution/Addition algorithm for partitioning the covariate space in prediction, *Bioinformatics*, 26 (10), 1357–1363.
162. Ori M. Stitelman and Mark J. van der Laan (2010), Collaborative Targeted Maximum Likelihood For Time To Event Data, (March 2010). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 260, <http://www.bepress.com/ucbbiostat/paper260>, *The International Journal of Biostatistics*: Vol. 6 : Iss. 1, Article 21.
163. Moore, K., van der Laan, M. J. (2009), Increasing Power in Randomized Trials with Right Censored Outcomes Through Covariate Adjustment. *Journal of Biopharmaceutical Statistics*, 19 (6): 1099-1131, 2009.
164. Sherri Rose and Mark J. van der Laan (2010), A targeted maximum likelihood estimator for two-stage designs *Int J Biostat* (2011) <http://www.bepress.com/ijb/vol7/iss1/17/>

165. ML Petersen, KE Porter, S Gruber, Y Wang, MJ van der Laan (2010), Diagnosing and responding to violations in the positivity assumption, *Stat Meth Med Res*, <http://smm.sagepub.com>
166. A. Chambaz and M. J. van der Laan (2011), Targeting The Optimal Design In Randomized Clinical Trials With Binary Outcomes And No Covariate Part I. U.C. Berkeley Division of Biostatistics Working Paper Series, Working Paper 258. *International Journal of Biostatistics*, number 1, pages 1-32, volume 7.
167. A. Chambaz and M. J. van der Laan (2011), Targeting The Optimal Design In Randomized Clinical Trials With Binary Outcomes And No Covariate Part II. U.C. Berkeley Division of Biostatistics Working Paper Series, Working Paper 258. *International Journal of Biostatistics*. number 1, pages 33- , volume 7.
168. Susan Gruber, Mark van der Laan (2010), A Targeted Maximum Likelihood Estimator of a Causal Effect on a Bounded Continuous Outcome, *The International Journal of Biostatistics: Vol. 6 : Iss. 1, Article 26*.
169. Stitelman, Ori M.; Wester, C. William; De Gruttola, Victor; and van der Laan, Mark J. (2011), Targeted Maximum Likelihood Estimation of Effect Modification Parameters in Survival Analysis, *The International Journal of Biostatistics: Vol. 7: Iss. 1, Article 19*. DOI: 10.2202/1557-4679.1307 Available at: <http://www.bepress.com/ijb/vol7/iss1/19>
170. A. Chambaz and M. J. van der Laan (2011), Targeting The Optimal Design In Randomized Clinical Trials With Covariates, book chapter in *Targeted Learning, Estimation of Causal effects in Observational and Experimental Studies*, Springer Verlag, van der Laan, Rose (2011).
171. Catherine Tuglus and Mark J. van der Laan (2011), Repeated measures semiparametric regression using targeted maximum likelihood methodology with application to transcription factor activity discovery *Stat Appl Genet Molec Biol* (2011), <http://www.bepress.com/sagmb/vol10/iss1/art2/>
172. Catherine Tuglus and Mark J. van der Laan (2011), "Targeted Methods for Biomarker Discovery, the Search for a Standard" , Book chapter in *Targeted Learning: Estimation of Causal Effects in Experimental and Observational Studies* (2011), van der Laan, Rose, Springer New York.
173. H Wang, S Rose, MJ van der Laan (2010), Finding quantitative trait loci genes with collaborative targeted maximum likelihood learning, *Stat Probab Lett*. 2011 Jul 1;81(7):792-796.
174. McCulloch M1, Broffman M, van der Laan M, Hubbard A, Kushi L, Abrams DI, Gao J, Colford JM Jr. (2011), Colon Cancer Survival with Herbal Medicine & Vitamins in a Whole-Systems Approach: 10-Year Follow-Up Data Analyzed with Marginal Structural Models & Propensity Score Methods, *Integr Cancer Ther.*, Sep;10(3):240-59. doi: 10.1177/1534735411406539. Epub 2011 Sep 30.
175. McCulloch M1, Broffman M, van der Laan M, Hubbard A, Kushi L, Abrams DI, Gao J, Colford JM Jr. (2011), Lung Cancer Survival with Herbal Medicine & Vitamins in a Whole-Systems Approach: 10-Year Follow-Up Data Analyzed with Marginal Structural Models & Propensity Score Methods, *Integr Cancer Ther.* 2011 Sep;10(3):260-79. doi: 10.1177/1534735411406439. Epub 2011 Aug 8.
176. Michael Rosenblum, Mark van der Laan (2011), Optimizing randomized trial designs to distinguish which subpopulations benefit from treatment, *Biometrika*, 98: 845-860.

177. W. Zheng, M.J. van der Laan (2011), Asymptotic theory for cross-validated targeted maximum likelihood estimation, book chapter in "Targeted Learning: Estimation of Causal Effects in Experimental and Observational Studies", van der Laan, Rose (2011), Springer New York. Also available at <http://www.bepress.com/ucbbiostat/paper273/>
178. Moore KL, Neugebauer R, Valappil T, van der Laan MJ. Robust extraction of covariate information to improve estimation efficiency in randomized trials. *Statistics in Medicine*. 2011 Doi: 10.1002/sim.4301
179. I. Diaz Munoz, M.J. van der Laan (2011), Targeted Bayesian learning, book chapter in "Targeted Learning: Estimation of Causal Effects in Experimental and Observational Studies", van der Laan, Rose (2011), Springer New York. Also available at <http://www.bepress.com/ucbbiostat/paper270/>
180. EC Polley, S. Rose, MJ van der Laan (2011), Super learner in prediction, book chapter in "Targeted Learning: Estimation of Causal Effects in Experimental and Observational Studies", van der Laan, Rose (2011), Springer New York. Also available at <http://www.bepress.com/ucbbiostat/paper266/>
181. H. Wang, M.J. van der Laan (2011), Dimension Reduction with Gene Expression Data Using Targeted Variable Importance Measurement, *BMC Bioinformatics*, 12:312.
182. I. Diaz-Munoz, M.J. van der Laan (2012), Population intervention causal effects based on stochastic interventions, *Biometrics*. Jun;68(2):541-9. doi: 10.1111/j.1541-0420.2011.01685.x. Epub 2011 Oct 6.
183. K. Moore, R. Neugebauer, M.J. van der Laan, I.B. Tager (2012), Causal inference in epidemiological studies with strong confounding, *Stat Med*. 2012 Jun 15;31(13):1380-404. doi: 10.1002/sim.4469. Epub 2012 Feb 23.
184. Moore KL, Neugebauer R, Valappil T, van der Laan MJ. Robust extraction of covariate information to improve estimation efficiency in randomized trials. *Statistics in Medicine*. 2011 Doi: 10.1002/sim.4301.
185. Ori M. Stitelman, Victor De Gruttola, and Mark J. van der Laan (2012), "A General Implementation of TMLE for Longitudinal Data Applied to Causal Inference in Survival Analysis", *Int J Biostat*. 2012 Sep 18;8(1). pii: /j/ijb.2012.8.issue-1/1557-4679.1334/1557-4679.1334.xml. doi: 10.1515/1557-4679.1334. Also available at U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 281, <http://www.bepress.com/ucbbiostat/paper281>
186. Susan Gruber and Mark J. van der Laan (2012), "Targeted Minimum Loss Based Estimator that Outperforms a given Estimator", *Int J Biostat*. 2012 May 18;8(1). pii: /j/ijb.2012.8.issue-1/1557-4679.1332/1557-4679.1332.xml. doi: 10.1515/1557-4679.1332.(April 2011). Also available at: U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 280, <http://www.bepress.com/ucbbiostat/paper280>.
187. Kristin E. Porter, Susan Gruber, Mark J. van der Laan, and Jasjeet S. Sekhon (2011), "The Relative Performance of Targeted Maximum Likelihood Estimators", *Int J Biostat*. Jan 1, 2011; 7(1): Article 31. Published online Aug 17, 2011. doi: 10.2202/1557-4679.1308. Also available at: U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 279, <http://www.bepress.com/ucbbiostat/paper279>,

188. W. Zheng, M.J. van der Laan (2012), A Targeted Maximum Likelihood Estimator of Natural Direct Effects, *The International Journal of Biostatistics*. Volume 8, Issue 1, Pages 1:40, ISSN (Online) 1557-4679, DOI: 10.2202/1557-4679.1361, January 2012.
189. C.W. Wester, O. Stitelman, R. Marlink, H. Bussmann, V. DeGruttolas, M.J. van der Laan (2012), Effect Modification by Sex and Baseline CD4+ Cell Count Among Adults Receiving Combination Antiretroviral Therapy in Botswana: Results from a Clinical Trial, *AIDS Res Hum Retroviruses*. 2012 Sep;28(9):981-8. doi: 10.1089/AID.2011.0349. Epub 2012 Mar 23.
190. Anup Malani, Oliver Bembom and Mark van der Laan (2012), Accounting for Heterogeneous Treatment Effects in the FDA Approval Process, *Food Drug Law J*. 2012;67(1):23-50, i.
191. S. Gruber and M. J. van der Laan. Consistent causal effect estimation under dual misspecification and implications for confounder selection procedures. *Stat Methods Med Res*, 24(6):10031008, Dec 2015. PMID: 22368176.
192. Dan Rubin, Mark van der Laan (2012), A Case Study of Statistical Issues and Limitations in Personalized Medicine Research with Clinical Trials. *The International Journal of Biostatistics*. Volume 8, Issue 1, ISSN (Online) 1557-4679, DOI: 10.1515/1557-4679.1423, July 2012.
193. P Chaffee, MJ van der Laan (2012), Targeted maximum likelihood estimation for dynamic treatment regimes in sequential randomized controlled trials, *Int J Biostat*. 2012 Jun 22;8(1):Article 14. doi: 10.1515/1557-4679.1406. Also available at: <http://www.bepress.com/ucbbiostat/paper277/>
194. Antoine, Chambaz; Neuviel, Pierre; and van der Laan, Mark J (2012)., "Estimation of a Non-Parametric Variable Importance Measure of a Continuous Exposure", *Electron. J. Statist*. Volume 6 (2012), 1059-1099. Also available at: U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 292, <http://biostats.bepress.com/ucbbiostat/paper292A>.
195. M.J. van der Laan, M. Petersen (2012), Super Learning, book chapter in "Ensemble Machine Learning: Methods and Applications", Zhang, Cha, Ma, Yunqian (Eds.), Springer, New York.
196. Chaffee, PH and van der Laan MJ (2012), Discussion of Evaluation of Viable Dynamic Treatment Regimes in a Sequentially Randomized Trial of Advanced Prostate Cancer, *Journal of the American Statistical Association*, Volume 107, Issue 498, June 2012, 513-517.
197. Brooks J, van der Laan MJ. (2012), Targeted Maximum Likelihood Estimation for Prediction Calibration. *Int J Biostat*. 2012 Oct 31;8(1):30. doi: 10.1515/1557-4679.1385.
198. G. Geeven, M.J. van der Laan, M.C.M. de Gunst (2012), Comparison of targeted maximum likelihood and shrinkage estimators of parameters in gene networks, *Stat Appl Genet Mol Biol*. 2012 Sep 25;11(5):Article 2. doi: 10.1515/1544-6115.1728.
199. M. Rosenblum, M.J. van der Laan (2012), Response to Jane Paik's reader reaction to "Using regression models to analyze randomized trials: asymptotically valid hypothesis tests despite incorrectly specified models", *Biometrics*. doi: 10.1111/j.1541-0420.2012.01799.x

200. S. Gruber, M.J. van der Laan (2013), An application of targeted maximum likelihood estimation to the meta-analysis of safety data, *Biometrics*, Mar;69(1):254-62. doi: 10.1111/j.1541-0420.2012.01829.x. Epub 2013 Feb 4.
201. Legrand M, Pirracchio R, Rosa A, Petersen M, van der Laan M, Fabiani JN, Fernandez-Gerlinger MP, Podglajen I, Safran D, Cholley B, Mainardi J. Incidence, Risk Factors and Prediction of Post-operative Acute Kidney Injury Following Cardiac Surgery For Active Infective Endocarditis: an Observational Study. *Critical Care*. 2013 Oct 4;17(5):R220.
202. S. Lendle, Meenakshi S. Subbaraman, M.J. van der Laan (2012), Identification and efficient estimation of the natural direct effect among the untreated, *Biometrics* Volume 69, Issue 2, pages 310-317, June 2013.
203. S. Lendle, B. Fireman, M.J. van der Laan (2013), Targeted maximum likelihood estimation in safety analysis, *Journal of Clinical Epidemiology* Volume 66, Issue 8, Supplement, Pages S91-S98, August 2013.
204. M. Carone, M. Petersen, M.J. van der Laan (2012), Targeted minimum loss based estimation of a casual effect using interval censored time to event data. Chapter 8 of *Interval Censored Time To Event Data: Methods and Applications*. Chapman & Hall/CRC. Editors: Ding-Geng (Din) Chen, Jianguo Sun, Karl E. Peace.
205. T.J. Haight, M.J. van der Laan, I.B. Tager (2013), Direct effects of leisure-time physical activity on walking speed, *The Journal of Nutrition, Health & Aging*, <http://link.springer.com/article/10.1007>.
206. M.J. van der Laan, L. Balzer and M.L. Petersen (2012), Adaptive matching in randomized trials and observational studies, *Journal of Statistical Research*, Volume 46(2), 113–156.
207. * A. Chambaz and M. J. van der Laan. Inference in targeted group-sequential covariate-adjusted randomized clinical trials. *Scandinavian Journal of Statistics. Theory and Applications*, 41(1):104140, 2014. NIHMSID: NIHMS979199.
208. M.J. van der Laan, S. Gruber (2012), Targeted minimum loss based estimation of causal effects of multiple time point interventions, *Int J Biostat*. 2012;8(1). doi: 10.1515/1557-4679.1370.
209. M.J van der Laan (2012), Targeted estimation of nuisance parameters to obtain valid statistical inference, *Int J Biostat*. 2014;10(1):29-57. doi: 10.1515/ijb-2012-0038.
210. M. Subbaraman, S. Lendle, M. van der Laan, J. Ahern (2013), Cravings as a mediator and moderator of drinking outcomes in the COMBINE Study, *Addiction*. 2013 Oct;108(10):1737-44. doi: 10.1111/add.12238. Epub 2013 Jun 4.
211. S. Rose, M.J. van der Laan (2014), A double robust approach to causal effects in case-control studies, *Am J Epidemiol*. Volume 179, Number 6, 663-670, doi: 10.1093/aje/kwt318. Epub 2014 Jan 31.
212. M.J. van der Laan, M. Petersen, W. Zheng (2013), Estimating the effect of a community-based intervention with two communities, *Journal of Causal Inference* 2013, 1(1): 1-24.
213. M. Legrant, R. Pirracchio, A. Rosa, M. Petersen, M. van der Laan (2013), Incidence, risk factors and prediction of postoperative acute kidney injury following cardiac surgery for active infective endocarditis: an observational study, *Critical Care* 2013, 17:R220 doi:10.1186/cc13041.

214. Antoine Chambaz; Dominique Choudat; Catherine Huber; Jean-Claude Pairon; Mark J. van der Laan (2013), Analysis of the effect of occupational exposure to asbestos based on threshold regression modeling of case-control data, *Biostatistics* 2013; doi: 10.1093/biostatistics/kxt042
215. M. Petersen and J. Schwab and S. Gruber and N. Blaser and M. Schomaker and M.J. van der Laan (2013), Targeted Minimum Loss Based Estimation of Marginal Structural Working Models, *Journal of Causal Inference*. Volume 2, Issue 2, Pages 147-185, ISSN (Online) 2193-3685, ISSN (Print) 2193-3677, DOI: 10.1515/jci-2013-0007, June 2014. Also available at: <http://biostats.bepress.com/ucbbiostat/paper312/> PMID: PMC4405134.
216. M. Schnitzer, E. Moodie, M.J. van der Laan, R. Platt, M. Klein (2014), "Modeling the impact of hepatitis C viral clearance on end-stage liver disease in an HIV co-infected cohort with Targeted Maximum Likelihood Estimation," *Biometrics*, (70)(1), 144–152, Also available at: <http://biostats.bepress.com/ucbbiostat/paper304> PMID: 24571372.
217. Díaz, Ivan and van der Laan, Mark J. (2013), "Sensitivity Analysis for Causal Inference Under Unmeasured Confounding and Measurement Error Problems", *Int J Biostat*. 2013 Nov 19;9(2):149-60. doi: 10.1515/ijb-2013-0004. Also available at: U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 303. <http://biostats.bepress.com/ucbbiostat/paper303>.
218. Sapp, Stephanie; van der Laan, Mark J.; and Page (2014), Kimberly, "Targeted Estimation of Variable Importance Measures with Interval-Censored Outcomes", *The International Journal of Biostatistics*. Volume 10, Issue 1, Pages 77-97, ISSN (Online) 1557-4679, ISSN (Print) 2194-573X, DOI: 10.1515/ijb-2013-0009, March 2014. Also available at: U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 307. <http://biostats.bepress.com/ucbbiostat/paper307>.
219. I. Diaz and M. J. van der Laan (2013), Assessing the causal effect of policies: An approach based on stochastic interventions, *Int J Biostat*. 2013 Nov 19;9(2):161-74. doi: 10.1515/ijb-2013-0014.
220. Maya Petersen, Mark van der Laan (2013), Causal Models and Learning from Data: Integrating Causal Modeling and Statistical Estimation in the Practice of Epidemiology, *Epidemiology*. 2014 May;25(3):418-26. doi: 10.1097/EDE.0000000000000078. PMID: 24713881 (May 2014)
221. R. Starmans and M.J. van der Laan (2013), Inferentiele statistiek versus machine learning: een aanzet tot verzoening, *Stator* **42**, 1–4.
222. J. Brooks, M. J. van der Laan, D.E. Singer, A. S. Go (2013), Targeted Minimum Loss-Based Estimation of Causal Effects in Right-Censored Survival Data with Time-Dependent Covariates: Warfarin, Stroke, and Death in Atrial Fibrillation, *Journal of Causal Inference*. Volume 1, Issue 2, Pages 235-254, ISSN (Online) 2193-3685, ISSN (Print) 2193-3677, DOI: 10.1515/jci-2013-0001, November 2013.
223. R. Neugebauer, J. Schmittdiel, M.J. van der Laan (2014), Targeted Learning in real-world comparative effectiveness research with time-varying interventions, *Statistics in Medicine*, Volume 33, Issue 14, pages 2480-2520, 30 June 2014. PMID:24535915.
224. S. Rose, M.J. van der Laan (2014), Response to Invited commentary "Some advantages of the relative excess risk due to interaction", by T.J. VanderWeele and S. VanSteelandt, *American Journal of Epidemiology*, Volume 179, Number 6, page 672.

225. J. Brooks, D.E. Singer, A. S. Go, M.J. van der Laan (2014), Book chapter A roadmap for causal inference in safety analysis in *Quantitative Evaluation of Safety in Drug Development: Design, Analysis and Reporting*, Chapman and Hall/CRC, editors Q. Jang, A. Xia.
226. M.J. van der Laan, R.J.C.M. Starmans (2014), Review article: Entering the era of data science: Targeted Learning and the Integration of Statistics and Computational Data Analysis, *Advances in Statistics Volume 2014* (2014), Article ID 502678, 19 pages <http://dx.doi.org/10.1155/2014/502678>
227. M.J. van der Laan (2014), Targeted Learning: From MLE to TMLE, book chapter in *COPSS 50th Anniversary Volume, Past, Present and Future of Statistical Science*, CRC Press, Taylor & Francis Group.
228.] M. J. van der Laan. Causal Inference for a Population of Causally Connected Units. *J Causal Inference*, 2(1):1374, Mar 2014. PMID: 26180755.
229. M. E. Schnitzer, M. J. van der Laan, E. E. Moodie, and R. W. Platt. Effect of breastfeeding on gastrointestinal infection in infants: A targeted maximum likelihood approach for clustered longitudinal data. *Ann Appl Stat*, 8(2):703725, Jun 2014. PMID: 25505499.
230. H. Wang, Z. Zhang, S. Rose, M.J. van der Laan (2014), A Novel Targeted Learning Method for Quantitative Trait Loci Mapping, *Genetics*, 198(4), 1369–76. Special highlighted article. PMID: 25258376.
231. Romain Pirracchio, Maya Petersen, Matthieu Resche Rigon, Sylvie Chevret, Didier Journois, Denis Safran, Mark van der Laan (2014), Mortality prediction in the ICU: can we do better? Results from the Super ICU Learner Algorithm (SICULA) projection, a population-based study, Epub, Nov 24, 2014, *Lancet Respiratory Medicine*, doi:10.1016/S2213-2600(14)70239-5. PMID 25466337.
232. Diaz, Ivan, van der Laan, M.J. (2013), Targeted Data Adaptive Estimation of the Causal Dose–Response Curve, *Journal of Causal Inference*, 1, (2), p. 171–192.
233. Miguel E, Camerer C, Casey K, Cohen J, Esterling KM, Gerber A, Glennerster R, Green DP, Humphreys M, Imbens G, Laitin D, Madon T, Nelson L, Nosek BA, Petersen M, Sedlmayr R, Simmons JP, Simonsohn U, Van der Laan M. Social science. Promoting transparency in social science research. *Science*. 2014; 343(6166): 30-154.
234. Brown D, Petersen M, van der Laan M, Costello S, Noth E, Hammond K, Cullen M, Eisen E. 0124 PM2.5 and Heart Disease in a Cohort of Aluminium Workers: An Application of Longitudinal Targeted Maximum Likelihood-based Estimation (TMLE), *Occup Environ Med*. 2014 Jun;71 Suppl 1:A14. doi: 10.1136/oemed-2014-102362.44.
235. A. M. Weber, M. J. van der Laan, and M. L. Petersen. Assumption Trade-Offs When Choosing Identification Strategies for Pre-Post Treatment Effect Estimation: An Illustration of a Community-Based Intervention in Madagascar. *J Causal Inference*, 3(1):109130, Mar 2015. PMID: 26097800.
236. M J van der Laan, A. R. Luedtke and I Daz. Discussion of Identification, Estimation and Approximation of Risk under Interventions that Depend on the Natural Value of Treatment Using Observational Data, by Jessica Young, Miguel Hernn, and James Robins. In: *Epidemiologic Methods* 3.1 (2014), pp. 2131

237. M. J. van der Laan and A. R. Luedtke (2015), Targeted learning of the mean outcome under an optimal dynamic treatment rule, *Journal of Causal Inference*, vol. 3, no. 1, pp. 61-95, 2015. PMID: PMC4517487. Also available at <http://biostats.bepress.com/ucbbiostat/paper329/>
238. L.B. Balzer, M.L. Petersen, M.J. van der Laan (2014), Adaptive pair-matching in randomized trials with unbiased and efficient effect estimation. *Statistics in Medicine*, 34(6):999-1011, 2015. PMID: 25421503 Also available at <http://biostats.bepress.com/ucbbiostat/paper320/>
239. R. Pirracchio, M. L. Petersen, and M. van der Laan. Improving propensity score estimators robustness to model misspecification using super learner. *Am. J. Epidemiol.*, 181(2):108119, Jan 2015. PMID: 25515168.
240. Zheng W. Chambaz A, van der Laan M. Group Sequential clinical trials with response-adaptive randomization. Chapter in "Modern Adaptive Randomized Clinical Trials: Statistical, Operational, and Regulatory Aspects", Edited by Alex Sverdlov (Forthcoming 2015 Chapman Hall/CRC, London, 2015.
241. A. R. Luedtke, M. Carone, and M. J. van der Laan. Discussion of Deductive derivation and turing- computerization of semiparametric efficient estimation by Frangakis et al. *Biometrics*, 71(4):875879, Dec 2015. PMID: 26243187.
242. M. L. Petersen and M. J. van der Laan. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*, 25(3):418426, May 2014. PMID: 24713881.
243. D. M. Brown, M. Petersen, S. Costello, E. M. Noth, K. Hammond, M. Cullen, M. v. Laan, and E. Eisen. Occupational Exposure to PM2.5 and Incidence of Ischemic Heart Disease: Longitudinal Targeted Minimum Loss-based Estimation. *Epidemiology*, 26(6):806814, Nov 2015. PMID: PMC26079662. appear in *Epidemiology*.
244. Petersen M, Schwab J, Geng E, van der Laan M. Comparing the effectiveness of dynamic regimes with and without marginal structural models. In: *Dynamic Treatment Regimes in Practice: Planning Trials and Analyzing Data for Personalized Medicine*. (Moodie E and Kosorok M, eds.). 2016, ASA-SIAM, Philadelphia, PA. <http://bookstore.siam.org/sa21/>
245. Stephanie Sapp, Mark J. van der Laan, and John Canny (2014). Subsemble: an ensemble method for combining subset-specific algorithms. *Journal of Applied Statistics*, 41(6):1247, 1259. PMID: 24778462.
246. Hubbard, S K-Pajouh, van der Laan (2016), Statistical Inference for Data Adaptive Target Parameters, *Int J Biostat.* 2016 May 1;12(1):3-19. doi: 10.1515/ijb-2015-0013.
247. M.J van der Laan (2015), *Observational Studies* 1 (2015), 220-222, Comment on "Observational Studies" by Dr. W.G. Cochran (1972).
248. M. J. van der Laan, A. R. Luedtke, and I Díaz, Discussion of Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data, by Jessica Young, Miguel Hernán, and James Robins, *Epidemiologic Methods*, vol. 3, no. 1, pp. 21-31, 2014.
249. M. L. Petersen, E. LeDell, J. Schwab, V. Sarovar, R. Gross, N. Reynolds, J. E. Haberer, K. Goggin, C. Golin, J. Arnsten, M. I. Rosen, R. H. Remien, D. Etoori, I. B. Wilson, J. M. Simoni, J. A. Erlen, M. J. van der Laan, H. Liu, and D. R. Bangsberg. Super Learner Analysis of Electronic Adherence Data Improves Viral Prediction and

- May Provide Strategies for Selective HIV RNA Monitoring. *J. Acquir. Immune Defic. Syndr.*, 69(1):109118, May 2015. PMID: 25942462.
250. Geng E van der Laan M. Petersen M, Schwab J. Illustration of analytic methods based on the ADAPT-R Trial. ASA-SIAM, Philadelphia, 2015.
 251. J. Ahern, D. Kasarek, A. R. Luedtke, T. A. Bruckner, and M. J. van der Laan (2016), Racial/ethnic differences in the role of childhood adversities for mental health disorders among a nationally representative sample of adolescents, *Epidemiology*, vol. 27, no. 5, pp. 697-704 2016.
 252. L. Balzer, M. Petersen, and M. van der Laan. Adaptive pair-matching in randomized trials with unbiased and efficient effect estimation. *Stat Med*, 34(6):9991011, Mar 2015. PMID: PMC4318754.
 253. Lendle, Samuel D and Fireman, Bruce and van der Laan, Mark J (2015), Balancing Score Adjusted Targeted Minimum Loss-based Estimation, *Journal of Causal Inference*. Volume 3, Issue 2, Pages 139-155, ISSN (Online) 2193-3685, ISSN (Print) 2193-3677, DOI: 10.1515/jci-2012-0012, January 2015.
 254. Le Dell, Petersen M, van der Laan M (2015), Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates, *Electronic Journal of Statistics*, Volume 9, Number 1, 1583–1607. PMID: PMC4533123.
 255. L. Balzer, J. Ahern, S. Galea, and M. van der Laan. Estimating Effects with Rare Outcomes and High Dimensional Covariates: Knowledge is Power. *Epidemiol Methods*, 5(1):118, Dec 2016. PMID: 28529839. Technical report available at <http://biostats.bepress.com/ucbbiostat/paper310/>
 256. L. Balzer, M. van der Laan, and M. Petersen. Adaptive pre-specification in randomized trials with and without pair-matching. *Stat Med*, 35(25):45284545, Nov 2016. PMID: PMC5084457.
 257. S. D. Lendle, J. Schwab, M. L. Petersen, and M. J. van der Laan. ltmle: An R package implementing targeted minimum loss-based estimation for longitudinal data. *Journal of Statistical Software*, 81(1):121, 2017.
 258. R. Pirracchio, J. Yue, G.T. Manley, M.J. van der Laan, A.E. Hubbard (2016), Collaborative targeted maximum likelihood estimation for variable importance measure: Illustration for functional outcome prediction in mild traumatic brain injuries. *Statistical Methods in Medical Research*, page 0962280215627335.
 259. A. R. Luedtke and M. J. van der Laan (2016), Comment, *Journal of the American Statistical Association*, vol. 111, no. 516, pp. 1526-1530.
 260. A. R. Luedtke and M. J. van der Laan (2016), Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy, *Annals of Statistics*, vol. 44, no. 2, pp. 713-742, 2016. NIHMSID: NIHMS979186.
 261. A. Luedtke and M. van der Laan. Optimal Individualized Treatments in Resource-Limited Settings. *Int J Biostat*, 12(1):283303, 05 2016. PMID: 27227725. NIHMSID: NIHMS979194.
 262. A. R. Luedtke and M. J. van der Laan (2016), Super-learning of an optimal dynamic treatment rule, *International Journal of Biostatistics*, vol. 12, no. 1, pp. 305-332, 2016. also available at <http://biostats.bepress.com/ucbbiostat/paper326/> PMID: 27227726. NIHMSID: NIHMS979189.

263. D. Benkeser and M. van der Laan. The highly adaptive lasso estimator. In 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pages 689696, Oct 2016. PMID: PMC5662030. NIHMSID: NIHMS979192.
264. † Luedtke AR, van der Laan MJ. Parametric-rate inference for one-sided differentiable parameters. *J Am Stat Assoc.* 2018;113(522):780-788. doi: 10.1080/01621459.2017.1285777. Epub 2017 Feb 28. PMID: 30078921; PMID: PMC6075853.
265. D. Benkeser, C. Ju, S. Lendle, and M. van der Laan. Online cross-validation-based ensemble learning. *Stat Med*, 37(2):249260, Jan 2018. PMID: PMC5671383.
266. R. Neugebauer, J. A. Schmittdiel, A. S. Adams, R. W. Grant, and M. J. van der Laan. Identification of the joint effect of a dynamic treatment intervention and a stochastic monitoring intervention under the no direct effect assumption. *J Causal Inference*, 5(1), 03 2017. PMID: 29238650.
267. R. Neugebauer, J. A. Schmittdiel, and M. J. van der Laan. A Case Study of the Impact of Data-Adaptive Versus Model-Based Estimation of the Propensity Scores on Causal Inferences from Three Inverse Probability Weighting Estimators. *Int J Biostat*, 12(1):131155, 05 2016. PMID: 27227720. NIHMSID: NIHMS979154.
268. O. Sofrygin, M.J. van der Laan, R. Neugebauer (2016), *simcausal R Package: Conducting Transparent and Reproducible Simulation Studies of Causal Effect Estimation with Complex Longitudinal Data.* *Journal of Statistical Software.*
269. W. Zheng, L. Balzer, M. van der Laan, and M. Petersen. Constrained binary classification using ensemble learning: an application to cost-efficient targeted PrEP strategies. *Stat Med*, 37(2):261279, Jan 2018. PMID: PMC5701877. PMID: 28384841
270. W. Zheng, M. Petersen, M.J. van der Laan (2016), Doubly Robust and Efficient Estimation of Marginal Structural Models for the Hazard Function. *Int J Biostat.* 2016 May 1;12(1):233-52. doi: 10.1515/ijb-2015-0036. PMID: 27227723
271. M. J. van der Laan and S. Gruber. One-Step Targeted Minimum Loss-based Estimation Based on Universal Least Favorable One-Dimensional Submodels. *Int J Biostat*, 12(1):351378, 05 2016. PMID: PMC4912007.
272. W. Zheng and M. van der Laan. Longitudinal Mediation Analysis with Time-varying Mediators and Exposures, with Application to Survival Outcomes. *J Causal Inference*, 5(2), 09 2017. PMID: PMC5788048.
273. A. Chambaz, W. Zheng, and M. J. van der Laan. Targeted sequential design for targeted learning inference of the optimal treatment rule and its mean reward. *Ann. Statist.*, 45(6):25372564, 12 2017. PMID: PMC5794253.
274. A. Chambaz, M.J. van der Laan, W. Zheng (2018), Targeting a simple statistical bandit problem, Chapter in *Targeted Learning in Data Science* Edited by Mark van der Laan and Sherri Rose.
275. † W. Zheng and M.J. van der Laan (2018), Longitudinal Mediation Analysis with Time-varying Mediators and Exposures, Chapter in *Targeted Learning in Data Science* Edited by Mark van der Laan and Sherri Rose
276. L.B. Balzer, M.L. Petersen, M.J. van der Laan (2016), Targeted estimation and inference for the sample average treatment effect in trials with and without pair-matching; SEARCH Collaboration. *Stat Med.* 2016 Sep 20;35(21):3717-32. doi: 10.1002/sim.6965. Epub 2016 Apr 18. PMID:27087478.

277. L.B. Balzer, M.L. Petersen, M.J. van der Laan (2016), Adaptive pre-specification in randomized trials with and without pair-matching: SEARCH Collaboration. *Stat Med.* 2016 Nov 10;35(25):4528-4545. doi:10.1002/sim.7023. Epub 2016 Jul 19. PMID:27436797.
278. † Petersen M, Balzer L, Kwarsiima D, Sang N, Chamie G, Ayieko J, Kabami J, Owaraganise A, Liegler T, Mwangwa F, Kadede K, Jain V, Plenty A, Brown L, Lavoy G, Schwab J, Black D, van der Laan M, Bukusi EA, Cohen CR, Clark TD, Charlebois E, Kanya M, Havlir D. Association of Implementation of a Universal Testing and Treatment Intervention With HIV Diagnosis, Receipt of Antiretroviral Therapy, and Viral Suppression in East Africa. *JAMA.* 2017 Jun 6;317(21):2196-2206. doi: 10.1001/jama.2017.5705. PMID: 28586888; PMCID: PMC5734234.
279. E. LeDell, M. van der Laan, and M. Petersen. AUC-Maximizing Ensembles through Metalearning. *Int J Biostat*, 12(1):203218, 05 2016. PMCID: PMC4912128.
280. B. A. Goldstein, E. C. Polley, F. B. Briggs, M. J. van der Laan, and A. Hubbard. Testing the Relative Performance of Data Adaptive Prediction Algorithms: A Generalized Test of Conditional Risk Differences. *Int J Biostat*, 12(1):117129, 05 2016. PMID: 26529567.
281. L. Tran, C. Yiannoutsos, B. Musick, K. Wools-Kaloustian, A. Siika, S. Kimaiyo, M.J. van der Laan, and M. Petersen. Evaluating the Impact of a HIV Low-Risk Express Care Task-Shifting Program: A Case Study of the Targeted Learning Roadmap. *Epidemiol Methods*, 5(1):6991, Dec 2016. PMCID: PMC5520542.
282. W. Zheng, M. Petersen, and M.J. van der Laan. Doubly Robust and Efficient Estimation of Marginal Structural Models for the Hazard Function. *Int J Biostat*, 12(1):233252, 05 2016. PMCID: PMC4912008.
283. I. Diaz, M.J. van der Laan (2017), Doubly Robust Inference for Targeted Minimum Loss Based Estimation in Randomized Trials with Missing Outcome Data, *Statistics in Medicine*, Article ID: SIM7389 Article DOI: 10.1002/sim.7389, Internal Article ID: 14321211.
284. D. Benkeser, P. Gilbert, M.J. van der Laan and M. Carone (2017), Doubly robust nonparametric inference on the average treatment effect, *Biometrika*, , asx053, <https://doi.org/10.1093/biomet/asx053>
285. M.J. van der Laan, D. Benkeser, and O. Sofrygin (2017), Targeted minimum loss-based estimation, *Wiley StatsRef-Statistics Reference Online*.
286. K. E. Rudolph and M. J. van der Laan. Robust estimation of encouragement-design intervention effects transported across sites. *J R Stat Soc Series B Stat Methodol*, 79(5):15091525, Nov 2017. PMID: 29375249
287. O. Sofrygin, R. Neugebauer, M.J. van der Laan (2017). *simcausal* R Package: Conducting Transparent and Reproducible Simulation Studies of Causal Effect Estimation with Complex Longitudinal Data. *Journal of Statistical Software*, 81(2), pp. 147. doi: 10.18637/jss.v081.i02.
288. C. Ju, R. Wyss, J.M. Franklin, S. Schneeweiss, J. Haggstrom, and M.J. van der Laan. Collaborative- controlled LASSO for constructing propensity score-based estimators in high-dimensional data. *Stat Methods Med Res*, page 962280217744588, Jan 2017. PMID: 29226777. NIHMSID: NIHMS979175.

289. † O. Sofrygin and M. van der Laan. Semi-Parametric Estimation and Inference for the Mean Outcome of the Single Time-Point Intervention in a Causally Connected Population. *J Causal Inference*, 5(1), 03 2017. PMID: 5650205.
290. † Rudolph KE, Sofrygin O, Zheng W, van der Laan MJ. Robust and Flexible Estimation of Stochastic Mediation Effects: A Proposed Method and Example in a Randomized Trial Setting. *Epidemiology Methods*. 2018;7(1):20170007. doi: 10.1515/em-2017-0007. Epub 2017 Dec 13. PMID: 34026421; PMCID: PMC8136358.
291. M.J. van der Laan (2017), Targeted Learning: The Link from Statistics to Data Science, *Stator*, page 12-16, 4.
292. † Price BL, Gilbert PB, van der Laan MJ. Estimation of the optimal surrogate based on a randomized trial. *Biometrics*. 2018 Dec;74(4):1271-1281. doi: 10.1111/biom.12879. Epub 2018 Apr 27. PMID: 29701875; PMCID: PMC6393111.
293. † Ju C, Schwab J, van der Laan MJ. On adaptive propensity score truncation in causal inference. *Stat Methods Med Res*. 2019 Jun;28(6):1741-1760. doi: 10.1177/0962280218774817. Epub 2018 Jul 11. PMID: 29991330.
294. B. Toth, M.J. van der Laan (2018), Targeted learning of optimal individualized treatment rules under cost constraints. Chapter in *Biopharmaceutical Applied Statistics Symposium*, Springer Nature Singapore Pte Ltd, K. E. Peace et al. (eds.), *Biopharmaceutical Applied Statistics Symposium*, ICSA Book Series in Statistics, <https://doi.org/10.1007/978-981-10-7820-01>.
295. L. Balzer, W. Zheng, M.J. van der Laan, M. Petersen (2018), Statistical Methods in Medical Research. A new approach to hierarchical data analysis: Targeted maximum likelihood estimation for the causal effect of a cluster-level exposure, <http://journals.sagepub.com/eprint/7Gk45JhbBGIBa5zxTvBx/full>.
296. † (update) Davies, Molly Margaret and van der Laan, Mark J.. "Optimal Spatial Prediction Using Ensemble Machine Learning" *The International Journal of Biostatistics*, vol. 12, no. 1, 2016, pp. 179-201. <https://doi.org/10.1515/ijb-2014-0060>. <http://biostats.bepress.com/ucbbiostat/paper305>.
297. Petersen, M.; LeDell, E.; Sarovar, V.; Schwab, J.; Etoori, D.; Gross, R.; Wilson, I.; Golin, C.; Reynolds, N.; Remien, R.; Goggin, K.; Simoni, J.; Rosen, M.; Liu, H.; van der Laan, M.; Bangsberg, D. Classification of Virologic Failure using Electronic Adherence Data: An analysis of the MACH14 Consortium. To appear *JAIDS*.
298. † Ju C, Gruber S, Lendle SD, Chambaz A, Franklin JM, Wyss R, Schneeweiss S, van der Laan MJ. Scalable collaborative targeted learning for high-dimensional data. *Stat Methods Med Res*. 2019 Feb;28(2):532-554. doi: 10.1177/0962280217729845. Epub 2017 Sep 22. PMID: 28936917; PMCID: PMC6086775.
299. Lendle S, Schwab J, Petersen M, van der Laan M (2014). *ltnle*: An R Package Implementing Targeted Minimum Loss-based Estimation for Longitudinal Data, to appear in *Journal of Statistical Software*.
300. M. Carone, I. Diaz, M.J. van der Laan (2018), Higher order targeted minimum loss-based estimation, Chapter in *Targeted Learning in Data Science* Edited by Mark van der Laan and Sherri Rose, Springer.
301. L. B. Balzer, M. J. van der Laan, and M. L. Petersen . *Data-Adaptive Estimation in Cluster Randomized Trials*, Chapter in *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*, 195215. Springer International Publishing, Cham, 2018.

302. A. Hubbard, M.J. van der Laan (2016) Mining with inference: Data-adaptive target parameters. In P. Buhlmann, P. Drineas, M. Kane, and M. van der Laan, editors, Handbook of Big Data. Chapman & Hall/CRC, 2016.
303. L. Balzer, M. Petersen, and M.J. van der Laan. Tutorial for causal inference. In P. Buhlmann, P. Drineas, M. Kane, and M. van der Laan, editors, Handbook of Big Data. Chapman & Hall/CRC, 2016.
304. A. Chambaz, W. Zheng, and M. J. van der Laan. Targeting a Simple Statistical Bandit Problem, Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 437-451. Springer International Publishing, Cham, 2018.
305. O. Sofrygin and M. J. van der Laan (2018). Causal Inference in Longitudinal Network-Dependent Data, chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 349-371. Springer International Publishing, Cham, 2018.
306. M. van der Laan. A Generally Efficient Targeted Minimum Loss Based Estimator based on the Highly Adaptive Lasso. Int J Biostat, 13(2), Oct 2017. PMID: 29023235.
307. M. J. van der Laan. A Generally Efficient HAL-TMLE, chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 95-102. Springer International Publishing, Cham, 2018. NIHMSID: NIHMS979184.
308. M. J. van der Laan (2018), HAL Estimator of the Efficient Influence Curve, Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, 103-123. Springer International Publishing, Cham, 2018.
309. M. J. van der Laan and D. Benkeser (2018). Highly Adaptive Lasso (HAL), Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 77-94. Springer International Publishing, Cham, 2018.
310. M. J. van der Laan, W. Cai, and S. Gruber. One-Step TMLE (2018), Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 51-75. Springer International Publishing, Cham, 2018.
311. M. J. van der Laan., A. Chambaz, and C. Ju. C-TMLE for Continuous Tuning (2018), Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 143-161. Springer International Publishing, Cham, 2018.
312. M. J. van der Laan, A. Chambaz, and S. Lendle (2018). Online Targeted Learning for Time Series, Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 317-346. Springer International Publishing, Cham, 2018.
313. M.J. van der Laan and D. Benkeser (2018). Online Super Learning, Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 303-315. Springer International Publishing, Cham, 2018.
314. W. Zheng and M. J. van der Laan (2018). Mediation Analysis with Time-Varying Mediators and Exposures, Chapter in Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies, pages 277-299. Springer International Publishing, Cham, 2018.
315. † (Update) Carone M, Luedtke AR, van der Laan MJ. Toward computerized efficient estimation in infinite-dimensional models. J Am Stat Assoc. 2019;114(527):1174-1190. doi: 10.1080/01621459.2018.1482752. Epub 2018 Sep 13. PMID: 32405108; PMCID: PMC7219981.

316. † Ju, Cheng, Aurlien Bibaut, and Mark van der Laan. "The relative performance of ensemble methods with deep convolutional neural networks for image classification." *Journal of Applied Statistics* (2018): 1-19.
317. * (update from submitted to accepted) Marginal Structural Models with Counterfactual Effect Modifiers. Zheng W, Luo Z, van der Laan MJ. *Int J Biostat*. 2018 Jun 8;14(1):/j/ijb.2018.14.issue-1/ijb-2018-0039/ijb-2018-0039.xml. doi: 10.1515/ijb-2018-0039. PMID: 29883322
318. Wyss, Richard, Sebastian Schneeweiss, Mark van der Laan, Samuel D. Lendle, Cheng Ju, and Jessica M. Franklin. "Using super learner prediction modeling to improve high-dimensional propensity score estimation." *Epidemiology* 29, no. 1 (2018): 96-106.
319. * H. Qiu, A. Luedtke, M.J. van der Laan (2019), Comment on Entropy Learning for Dynamic Treatment Regimes, by B. Jiang and R. Song et al. *Statistica Sinica*, Volume 29, number 4, 1666-1678, <http://www.stat.sinica.edu.tw/statistica/>.
320. * Sofrygin O, Zhu Z, Schmittdiel JA, Adams AS, Grant RW, van der Laan MJ, Neugebauer R. Targeted learning with daily EHR data. *Stat Med*. 2019 Jul 20;38(16):3073-3090. doi: 10.1002/sim.8164. Epub 2019 Apr 25. PMID: 31025411.
321. * S. Gruber, M.J. van der Laan (2019), Comment on "Automated Versus Do-It-Yourself Methods for Causal Inference: Lessons Learned from a Data Analysis Competition, *Statistical Science* 2019, Vol. 34, No. 1, 82-85, <http://dx.doi.org/10.1214/18-STS689>.
322. * M.J. van der Laan, D. Benkeser, W. Cai (2019), Causal inference based on undersmoothing highly adaptive lasso, *Proceedings of https://why19.causalai.net/papers/laan-why19.pdf*
323. * Ju C, Benkeser D, van der Laan MJ. Robust inference on the average treatment effect using the outcome highly adaptive lasso. *Biometrics*. 2020 Mar;76(1):109-118. doi: 10.1111/biom.13121. Epub 2019 Oct 30. PMID: 31350906.
324. * Schnitzer ME, Sango J, Ferreira Guerra S, van der Laan MJ. Data-adaptive longitudinal model selection in causal inference with collaborative targeted minimum loss-based estimation. *Biometrics*. 2020 Mar;76(1):145-157. doi: 10.1111/biom.13135. Epub 2019 Nov 6. PMID: 31397506.
325. * Cai W, van der Laan MJ. One-step targeted maximum likelihood estimation for time-to-event outcomes. *Biometrics*. 2020 Sep;76(3):722-733. doi: 10.1111/biom.13172. Epub 2019 Nov 28. PMID: 31729005.
326. * Balzer LB, Ayieko J, Kwarisiima D, Chamie G, Charlebois ED, Schwab J, van der Laan MJ, Kanya MR, Havlir DV, Petersen ML. Far from MCAR: Obtaining Population-level Estimates of HIV Viral Suppression. *Epidemiology*. 2020 Sep;31(5):620-627. doi: 10.1097/EDE.0000000000001215. Erratum in: *Epidemiology*. 2021 Sep 1;32(5):e25. PMID: 32452912; PMCID: PMC8105880.
327. * Kreif N, Sofrygin O, Schmittdiel JA, Adams AS, Grant RW, Zhu Z, van der Laan MJ, Neugebauer R. Exploiting nonsystematic covariate monitoring to broaden the scope of evidence about the causal effects of adaptive treatment strategies. *Biometrics*. 2021 Mar;77(1):329-342. doi: 10.1111/biom.13271. Epub 2020 Apr 27. PMID: 32297311.

328. * A machine learning-based approach for estimating and testing associations with multivariate outcomes (2020) ,David Benkeser, Andrew Mertens, John M Colford, Alan Hubbard, Benjamin F Arnold, Aryeh Stein, Mark J van der Laan. The International Journal of Biostatistics 1(ahead-of-print), De Gruyter, 2020/?8/?12.
329. * Alex Luedtke, Aurlien Bibaut, Mark van der Laan (2021), "Correction note: "Statistical inference for the mean outcome under a possibly nonunique optimal treatment rule"" Annals of Statistics 2021, Vol. 49, No. 4, 2429-2430.
330. * Li H, Rosete S, Coyle J, Phillips RV, Hejazi NS, Malenica I, Arnold BF, Benjamin-Chung J, Mertens A, Colford JM Jr, van der Laan MJ, Hubbard AE. Evaluating the robustness of targeted maximum likelihood estimators via realistic simulations in nutrition intervention trials. *Stat Med.* 2022 May 30;41(12):2132-2165. doi: 10.1002/sim.9348. Epub 2022 Feb 16. PMID: 35172378.
331. * VanderWeele TJ, van der Laan MJ, Kessler RC. The Authors Respond. *Epidemiology.* 2020 Jul;31(4):e34. doi: 10.1097/EDE.0000000000001190. PMID: 32483070.
332. * VanderWeele TJ, Luedtke AR, van der Laan MJ, Kessler RC. The Authors Respond. *Epidemiology.* 2020 May;31(3):e31. doi: 10.1097/EDE.0000000000001157. PMID: 31880639.
333. * Gennatas ED, Friedman JH, Ungar LH, Pirracchio R, Eaton E, Reichmann LG, Interian Y, Luna JM, Simone CB 2nd, Auerbach A, Delgado E, van der Laan MJ, Solberg TD, Valdes G. Expert-augmented machine learning. *Proc Natl Acad Sci U S A.* 2020 Mar 3;117(9):4571-4577. doi: 10.1073/pnas.1906831117. Epub 2020 Feb 18. PMID: 32071251; PMCID: PMC7060733.
334. * Valdes G, Interian Y, Gennatas E, Van der Laan M. The Conditional Super Learner. *IEEE Trans Pattern Anal Mach Intell.* 2021 Dec 1;PP. doi: 10.1109/TPAMI.2021.3131976. Epub ahead of print. PMID: 34851823.
335. * Bibaut A, Chambaz A, Dimakopoulou M, Kallus N, van der Laan M. Post-Contextual-Bandit Inference. *Adv Neural Inf Process Syst.* 2021 Dec;34:28548-28559. PMID: 35785105; PMCID: PMC9249103.
336. * Kellermayer R, van der Laan M. Validation of Predictive Models in Inflammatory Bowel Diseases. *Gastroenterology.* 2021 Jan;160(1):21-22. doi: 10.1053/j.gastro.2020.11.005. Epub 2020 Nov 12. PMID: 33188805.
337. * Rudolph KE, Levy J, van der Laan MJ. Transporting stochastic direct and indirect effects to new populations. *Biometrics.* 2021 Mar;77(1):197-211. doi: 10.1111/biom.13274. Epub 2020 May 4. PMID: 32277465; PMCID: PMC7664994.
338. * Rudolph KE, Sofrygin O, van der Laan MJ. Complier stochastic direct effects: identification and robust estimation. *J Am Stat Assoc.* 2021;116(535):1254-1264. doi: 10.1080/01621459.2019.1704292. Epub 2020 Jan 23. PMID: 34531623; PMCID: PMC8439556.
339. * Benkeser D, Petersen M, van der Laan MJ. Improved small-sample estimation of nonlinear cross-validated prediction metrics. *J Am Stat Assoc.* 2020;115(532):1917-1932. doi: 10.1080/01621459.2019.1668794. Epub 2019 Oct 21. PMID: 33716360; PMCID: PMC7954141.
340. * Andrade-Pacheco R, Rerolle F, Lemoine J, Hernandez L, Met A, Juziwelo L, Bibaut AF, van der Laan MJ, Arnold BF, Sturrock HJW. Finding hotspots: development of an adaptive spatial sampling approach. *Sci Rep.* 2020 Jul 2;10(1):10939. doi: 10.1038/s41598-020-67666-3. PMID: 32616757; PMCID: PMC7331748.

341. * VanderWeele TJ, Luedtke AR, van der Laan MJ, Kessler RC. Selecting Optimal Subgroups for Treatment Using Many Covariates. *Epidemiology*. 2019 May;30(3):334-341. doi: 10.1097/EDE.0000000000000991. PMID: 30789432; PMCID: PMC6456380.
342. * Balzer LB, van der Laan M, Ayieko J, Kamya M, Chamie G, Schwab J, Havlir DV, Petersen ML. Two-Stage TMLE to reduce bias and improve efficiency in cluster randomized trials. *Biostatistics*. 2021 Dec 23:kxab043. doi: 10.1093/biostatistics/kxab043. Epub ahead of print. PMID: 34939083.
343. * Rudolph KE, Daz I, Hejazi NS, van der Laan MJ, Luo SX, Shulman M, Campbell A, Rotrosen J, Nunes EV. Explaining differential effects of medication for opioid use disorder using a novel approach incorporating mediating variables. *Addiction*. 2021 Aug;116(8):2094-2103. doi: 10.1111/add.15377. Epub 2021 Jan 22. PMID: 33340181.
344. * Hejazi NS, van der Laan MJ, Janes HE, Gilbert PB, Benkeser DC. Efficient non-parametric inference on the effects of stochastic interventions under two-phase sampling, with applications to vaccine efficacy trials. *Biometrics*. 2021 Dec;77(4):1241-1253. doi: 10.1111/biom.13375. Epub 2020 Sep 28. PMID: 32949147; PMCID: PMC8016405.
345. * Miles CH, Petersen M, van der Laan MJ. Causal inference when counterfactuals depend on the proportion of all subjects exposed. *Biometrics*. 2019 Sep;75(3):768-777. doi: 10.1111/biom.13034. Epub 2019 Apr 3. PMID: 30714118; PMCID: PMC6679813.
346. * Luedtke AR, Carone M, van der Laan MJ. An omnibus non-parametric test of equality in distribution for unknown functions. *J R Stat Soc Series B Stat Methodol*. 2019 Feb;81(1):75-99. doi: 10.1111/rssb.12299. Epub 2018 Nov 2. PMID: 31024219; PMCID: PMC6476331.
347. * Boileau P, Qi NT, van der Laan MJ, Dudoit S, Leng N. A flexible approach for predictive biomarker discovery. *Biostatistics*. 2022 Jul 21:kxac029. doi: 10.1093/biostatistics/kxac029. Epub ahead of print. PMID: 35861622.
348. * Cai W, van der Laan M. Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator. *Int J Biostat*. 2020 Aug 10:/j/ijb.ahead-of-print/ijb-2017-0070/ijb-2017-0070.xml. doi: 10.1515/ijb-2017-0070. Epub ahead of print. PMID: 32772002.
349. * Hejazi NS, Rudolph KE, Van Der Laan MJ, Daz I. Nonparametric causal mediation analysis for stochastic interventional (in)direct effects. *Biostatistics*. 2022 Feb 1:kxac002. doi: 10.1093/biostatistics/kxac002. Epub ahead of print. PMID: 35102366.
350. * Colson KE, Rudolph KE, Zimmerman SC, Goin DE, Stuart EA, Laan Mv, Ahern J. Optimizing matching and analysis combinations for estimating causal effects. *Sci Rep*. 2016 Mar 16;6:23222. doi: 10.1038/srep23222. PMID: 26980444; PMCID: PMC4793248.
351. * Ertefaie A, Hejazi NS, van der Laan MJ. Nonparametric inverse-probability-weighted estimators based on the highly adaptive lasso. *Biometrics*. 2022 Jul 15. doi: 10.1111/biom.13719. Epub ahead of print. PMID: 35839293.
352. * Balzer LB, Havlir DV, Kamya MR, Chamie G, Charlebois ED, Clark TD, Koss CA, Kwarisiima D, Ayieko J, Sang N, Kabami J, Atukunda M, Jain V, Camlin CS,

- Cohen CR, Bukusi EA, Van Der Laan M, Petersen ML. Machine Learning to Identify Persons at High-Risk of Human Immunodeficiency Virus Acquisition in Rural Kenya and Uganda. *Clin Infect Dis*. 2020 Dec 3;71(9):2326-2333. doi: 10.1093/cid/ciz1096. PMID: 31697383; PMCID: PMC7904068.
353. * Havlir DV, Balzer LB, Charlebois ED, Clark TD, Kwarisiima D, Ayieko J, Kabami J, Sang N, Liegler T, Chamie G, Camlin CS, Jain V, Kadede K, Atukunda M, Ruel T, Shade SB, Ssemmondo E, Byonanebye DM, Mwangwa F, Owaraganise A, Olilo W, Black D, Snyman K, Burger R, Getahun M, Achando J, Awuonda B, Nakato H, Kironde J, Okiror S, Thirumurthy H, Koss C, Brown L, Marquez C, Schwab J, Lavoy G, Plenty A, Mugoma Wafula E, Omany P, Chen YH, Rooney JF, Bacon M, van der Laan M, Cohen CR, Bukusi E, Kanya MR, Petersen M. HIV Testing and Treatment with the Use of a Community Health Approach in Rural Africa. *N Engl J Med*. 2019 Jul 18;381(3):219-229. doi: 10.1056/NEJMoa1809866. PMID: 31314966; PMCID: PMC6748325.
354. * Westling T, van der Laan MJ, Carone M. Correcting an estimator of a multivariate monotone function with isotonic regression. *Electron J Stat*. 2020;14(2):3032-3069. doi: 10.1214/20-ejs1740. Epub 2020 Aug 15. PMID: 33981382; PMCID: PMC8112587.
355. * Ju C, Combs M, Lendle SD, Franklin JM, Wyss R, Schneeweiss S, van der Laan MJ. Propensity score prediction for electronic healthcare databases using Super Learner and High-dimensional Propensity Score Methods. *J Appl Stat*. 2019;46(12):2216-2236. doi: 10.1080/02664763.2019.1582614. Epub 2019 Feb 22. PMID: 32843815; PMCID: PMC7444746.
356. * Tran L, Yiannoutsos C, Wools-Kaloustian K, Siika A, van der Laan M, Petersen M. Double Robust Efficient Estimators of Longitudinal Treatment Effects: Comparative Performance in Simulations and a Case Study. *Int J Biostat*. 2019 Feb 26;15(2):/j/ijb.2019.15.issue-2/ijb-2017-0054/ijb-2017-0054.xml. doi: 10.1515/ijb-2017-0054. PMID: 30811344; PMCID: PMC6710167.
357. * Gianfrancesco MA, Trupin L, Shiboski S, van der Laan M, Graf J, Imboden J, Yazdany J, Schmajuk G. Smoking Is Associated with Higher Disease Activity in Rheumatoid Arthritis: A Longitudinal Study Controlling for Time-varying Covariates. *J Rheumatol*. 2019 Apr;46(4):370-375. doi: 10.3899/jrheum.180262. Epub 2018 Dec 1. PMID: 30504507; PMCID: PMC6445683.
358. * Balzer LB, Zheng W, van der Laan MJ, Petersen ML. A new approach to hierarchical data analysis: Targeted maximum likelihood estimation for the causal effect of a cluster-level exposure. *Stat Methods Med Res*. 2019 Jun;28(6):1761-1780. doi: 10.1177/0962280218774936. Epub 2018 Jun 19. PMID: 29921160; PMCID: PMC6173669.
359. * van der Laan MJ, Benkeser D, Cai W. Efficient estimation of pathwise differentiable target parameters with the undersmoothed highly adaptive lasso. *Int J Biostat*. 2022 Jul 15. doi: 10.1515/ijb-2019-0092. Epub ahead of print. PMID: 35851449.
360. * Montoya LM, van der Laan MJ, Luedtke AR, Skeem JL, Coyle JR, Petersen ML. The optimal dynamic treatment rule superlearner: considerations, performance, and application to criminal justice interventions. *Int J Biostat*. 2022 Jun 16. doi: 10.1515/ijb-2020-0127. Epub ahead of print. PMID: 35708222.

361. * Montoya LM, van der Laan MJ, Skeem JL, Petersen ML. Estimators for the value of the optimal dynamic treatment rule with application to criminal justice interventions. *Int J Biostat.* 2022 Jun 6. doi: 10.1515/ijb-2020-0128. Epub ahead of print. PMID: 35659857.
362. * Phillips RV, van der Laan MJ. Discussion on "Adaptive enrichment designs with a continuous biomarker" by Nigel Stallard. *Biometrics.* 2022 Mar 25. doi: 10.1111/biom.13640. Epub ahead of print. PMID: 35332936.
363. * Gruber S, Phillips RV, Lee H, van der Laan MJ. Data-Adaptive Selection of the Propensity Score Truncation Level for Inverse Probability Weighted and Targeted Maximum Likelihood Estimators of Marginal Point Treatment Effects. *Am J Epidemiol.* 2022 May 4:kwac087. doi: 10.1093/aje/kwac087. Epub ahead of print. PMID: 35512316.
364. * David Benkeser, Weixin Cai, M.J. van der Laan (2020), A nonparametric super-efficient estimator of the average treatment effect, *Statist. Sci.* 35 (3), 484-495, (August 2020) DOI: 10.1214/19-STS735 Also discussion paper: Comment by Mireille Schnitzer, Increasing real world usage of TMLE, *Statist. Sci.* 35(3): 496-498 (August 2020). DOI: 10.1214/20-STS770
Rejoinder: A nonparametric super efficient estimator of the ATE, Benkeser, Cai, van der Laan (2020)
365. * Nonparametric efficient causal mediation with intermediate confounders I Daz, NS Hejazi, KE Rudolph, MJ van Der Laan - *Biometrika*, 2021
366. * cvCovEst: Cross-validated covariance matrix estimator selection and evaluation inR P Boileau, NS Hejazi, B Collica, MJ van der Laan - *Journal of Open Source Software*, 2021
367. * Continuous-time targeted minimum loss-based estimation of intervention-specific mean outcomes, to appear in *Annals of Statistics*, HC Rytgaard, TA Gerds, MJ van der Laan - arXiv preprint arXiv:2105.02088, 2021. file:///Users/laan/Downloads/AOS1906-006R3A0.pdf
368. * The current landscape in biostatistics of real-world data and evidence: causal inference frameworks for study design and analysis M Ho, M van der Laan, H Lee, J Chen, K Lee, Y Fang - *Statistics in Biopharmaceutical Research*, 2021
369. * The current landscape in biostatistics of real-world data and evidence: clinical study design and analysis J Chen, M Ho, K Lee, Y Song, Y Fang, BA Goldstein - *Statistics in Biopharmaceutical Research*, 2021
370. * Levy, Jonathan, van der Laan, Mark, Hubbard, Alan and Pirracchio, Romain. "A fundamental measure of treatment effect heterogeneity" *Journal of Causal Inference*, vol. 9, no. 1, 2021, pp. 83-108. <https://doi.org/10.1515/jci-2019-0003>
371. * Rachael V. Phillips and Mark J. van der Laans contribution to the Discussion of Assumption lean inference for generalized linear model parameters by Vansteelandt and Dukes RV Phillips, MJ van der Laan - *Journal of the Royal Statistical Society Series B*, 2022
372. * Risk minimization from adaptively collected data: Guarantees for supervised and policy learning A Bibaut, N Kallus, M Dimakopoulou, A Chambaz, M. J. van der Laan- *Advances in Neural Information Processing Systems*, 2021

373. * Generalized Policy Elimination: an efficient algorithm for Nonparametric Contextual Bandits A Bibaut, A Chambaz, M Laan - Conference on Uncertainty in Artificial Intelligence, 2020
374. * S. Rose, M.J. van der Laan (2022), Chapman & Hall/CRC, Handbook of Matching and Weighting Adjustments for Causal Inference, Why Machine Learning Cannot Ignore Maximum Likelihood Estimation. , ArXiv abs/2110.12112, to appear.

SUBMITTED/REVISED PUBLICATIONS

See www.bepress.com, working paper series, Division of Biostatistics, UC Berkeley.

375. A. Schuler, M.J. van der Laan (2022), The selective adaptive lasso, <https://arxiv.org/pdf/2205.10697.pdf>. To be submitted.
376. * van der Laan, Zeyi Wang, Lars van der Laan (2021), Higher Order Targeted Maximum Likelihood Estimation. <https://arxiv.org/pdf/2101.06290.pdf>, under revision.
377. * Comparative methods for the analysis of cluster randomized trials A Benitez, ML Petersen, MJ van der Laan, N Santos - arXiv preprint arXiv:2110.09633, 2021, submitted.
378. * Personalized Online Machine Learning I Malenica, RV Phillips, R Pirracchio, A Chambaz - arXiv preprint arXiv:2109.10452, 2021, under revision for Statistics in Medicine.
379. * Development of an ensemble machine learning prognostic model to predict 60-day risk of major adverse cardiac events in adults with chest pain CJ Kennedy, DG Mark, J Huang, MJ van der Laan - MedRxiv, 2021, to be submitted.
380. * Highly Adaptive Lasso, technical report, to be submitted. M van der Laan - University of California, Berkeley, 2020
381. * Robust Estimation of Data-Dependent Causal Effects based on Observing a Single Time-Series MJ van der Laan, I Malenica - arXiv preprint arXiv:1809.00734, 2018, to be submitted.

R-PACKAGES

382. * methyvim: Targeted, robust, and model-free differential methylation analysis in R NS Hejazi, RV Phillips, AE Hubbard, MJ van der Laan - F1000Research, 2018
383. * Causal Longitudinal Mediation: <https://github.com/zy-wang1/calm>.
384. * Highly adaptive lasso: <https://github.com/tlverse/hal9001>
385. *Causal mediation analysis: <https://github.com/tlverse/tmle3mediate>
386. *Causal effect of shift intervention: <https://github.com/tlverse/tmle3shift>
387. *Density estimation with highly adaptive lasso: <https://github.com/nhejazi/haldensify>
388. * Variable importance analysis: <https://github.com/nhejazi/biotmle>
389. *Super learner for covariance matrix estimation: <https://github.com/PhilBoileau/cvCovEst>
390. * Scalable adaptive lasso: Alejandro Schuler, in progress, workable version available.
391. * Super learning of optimal dynamic treatment rule: <https://github.com/tlverse/tmle3mopttx>
392. * Causal inference with time series: <https://github.com/imalenica/tstmle3>
393. * TMLE of mean outcome under estimated optimal rule: <https://github.com/imalenica/tmle3adapt>
394. Contributed to resampling multiple testing methodology implemented in (see book #274): <http://www.bioconductor.org/packages/release/bioc/html/multtest.html>

395. Contributed to variable importance TMLE implemented in (see #203): <http://cran.r-project.org/web/packages/tmle.npvi/index.html>
396. K. Pollard, M. van der Laan, R package 2.26.0 hopach, <http://www.bioconductor.org/packages/release/bioc/html/hopach.html>
397. Erin LeDell, Maya Petersen and Mark van der Laan (2012). cvAUC: Cross-Validated Area Under the ROC Curve Confidence Intervals. R package version 1.0-0. <http://CRAN.R-project.org/package=cvAUC>
398. Erin LeDell, Stephanie Sapp and Mark van der Laan (2014). subsemble: An Ensemble Method for Combining Subset-Specific Algorithm Fits. R package version 0.0.9. <http://CRAN.R-project.org/package=subsemble>
399. Eric Polley, Mark van der Laan, R package 2.0-15 SuperLearner, <http://cran.r-project.org/web/packages/SuperLearner/index.html>
400. S. Gruber, M. van der Laan, R package 1.2.0-4 tmle, <http://cran.r-project.org/web/packages/tmle/index.html>.
401. Schwab JS, Petersen ML, van der Laan MJ, with contributions from Gruber S. (2013), ltmle: Longitudinal Targeted Maximum Likelihood Estimation, (R package). cran.r-project.org/web/packages/ltmle/index.html
402. D. Benkeser and I. Malenica. cvma: Machine learning-based summary of association with multivariate outcomes. <https://github.com/benkeser/cvma>, 2017. R package version 0.1.0.
403. J. Coyle. tmle3: Extensible tmle framework. <https://github.com/tlverse/tmle3>, 2018. R package version 0.0.0.9.
404. J. Coyle and N. Hejazi. hal9001: A fast and scalable Highly Adaptive LASSO. <https://github.com/jeremyrcoyle/hal9001>, 2018. R package version 0.1.2.
405. J. Coyle and N. Hejazi. origami: A generalized framework for cross-validation in r. The Journal of Open Source Software, 3(21), January 2018.
406. J. R. Coyle, N. S. Hejazi, I. Malenica, and O. Sofrygin. sl3: modern super learning with pipelines. <https://github.com/jeremyrcoyle/sl3>, 2018. R package version 0.1.0.
407. LeDell E, Sapp S, and van der Laan M. subsemble: An ensemble method for combining subset-specific algorithm fits., 2014. <http://cran.r-project.org/>: CRAN.
408. C. Ju, S. Gruber, and M. van der Laan. ctmle: Collaborative targeted maximum likelihood estimation, 2017. R package version 0.1.1.
409. S. Lendle. Targetedlearning.jl. <https://lendle.github.io/TargetedLearning.jl/>, 2015.
410. I. Malenica and M. J. van der Laan. tstmle: Data-adaptive estimation and inference for causal effects with a single time series. <https://github.com/podTockom/tstmle>, 2017. R package version 0.1.0.
411. I. Malenica and W. Zhang. medltmle: Estimation for natural mediation effect in longitudinal data. <https://github.com/podTockom/medltmle>, 2017. R package version 0.1.0.
412. O. Sofrygin, F. Blaauw, A. Chambaz, and M. van der Laan. condensier: Non-parametric multivariate conditional density estimation with binned histograms. <https://github.com/osofr/condensier>, 2017. R package version 0.0.3.
413. O. Sofrygin and M. J. van der Laan. tmlenet: Targeted maximum likelihood estimation for networks. <https://github.com/osofr/tmlenet>, 2015. R package version 0.1.9.

414. O. Sofrygin, M. J. van der Laan, and R. Neugebauer. *simcausal* R Package: Conducting Transparent and Reproducible Simulation Studies of Causal Effect Estimation with Complex Longitudinal Data. *J Stat Softw*, 81, 2017. PMID: 29104515.
415. C. Zhang, M. J. van der Laan, O. Sofrygin, and J. Ahern. *tmleCommunity*: Targeted maximum likelihood estimation for hierarchical data. <https://github.com/chizhangucb/tmleCommunity>, 2017. R package version.

BOOKS

416. M.J. van der Laan (1996), *Efficient and Inefficient Estimation in Semiparametric Models*. *CWI-tract 114*, Centre for Mathematics and Computer Science, Amsterdam, the Netherlands.
Book reviews appeared in a French Journal and in “Short Book Reviews” International Statistical Institute, Editor Dr. A.M. Herzberg. *Zentrallblatt für Mathematik* 838/96.
417. M.J. van der Laan, J.M. Robins (2002), *Unified methods for Censored Longitudinal Data and Causality*, Springer Verlag. New York.
418. S. Dudoit, M.J. van der Laan, *Resampling Based Multiple Testing with Applications to Genomics*, Springer Series of Statistics (2007).
419. M.J. van der Laan, S. Rose (2011), *Targeted Learning, Causal Inference for Observational and Experimental Data*, Springer New York.
420. *Handbook on Big Data*, Chapman & Hall (2015), edited by Peter Buhlmann, Petros Drineas, Michael Kane, Mark van der Laan.
421. M.J. van der Laan, S. Rose (2018), *Targeted Learning in Data Science - Causal Inference for Complex Longitudinal Studies*, (Springer Series in Statistics) Springer, New York.
422. * M.J. van der Laan (2021-): *TMLE and the Highly Adaptive Lasso* (in progress).
423. * *Targeted Learning Handbook* (2022), Chapman and Hall, N. Jehazi, I Malenica, R. Phillips, J. Coyle, A. Hubbard, M.J. van der Laan. <https://tlverse.org/tlverse-handbook/tlverse-handbook.pdf>.
424. * *Modern Causal Inference* (2022) A. Schuler, M.J van der Laan, book in progress.

OTHER PUBLISHED WORK

425. * Bibaut, Aurlien F. and Mark J. van der Laan. Fast rates for empirical risk minimization with cadlag losses with bounded sectional variation norm. *arXiv: Statistics Theory* (2019).
426. * Mertens A, Benjamin-Chung J, Colford JM, et al. Child wasting and concurrent stunting in low- and middle-income countries. *medRxiv*. Published online June 11, 2020:2020.06.09.20126979. doi:10.1101/2020.06.09.20126979. <https://www.medrxiv.org/content/10.1101/2020.06.09.20127001v2>
427. *Mertens A, Benjamin-Chung J, Colford JM, et al. Causes and consequences of child growth failure in low- and middle-income countries. *medRxiv*. Published online June 11, 2020:2020.06.09.20127100. doi:10.1101/2020.06.09.20127100. <https://www.medrxiv.org/content/10.1101/2020.06.09.20126979v2>

428. * Benjamin-Chung J, Mertens A, Colford JM, et al. Early childhood linear growth failure in low- and middle-income countries. medRxiv. Published online June 11, 2020:2020.06.09.20127001. doi:10.1101/2020.06.09.20127001. <https://www.medrxiv.org/content/10.1101/2020.06.09.20127100v2>
429. M.J. van der Laan, S. Rose (2010). Statistics ready for a revolution: Next generation of statisticians must build tools for massive data sets. *Amstat News*, 399:38-39.
430. M.J. van der Laan, Why we need a statistical revolution, <http://www.stats.org/super-learning-and-the-revolution-in-knowledge/>
431. M.J. van der Laan (2013), Published Interview on blog "Musings on Iraq", by Joel Wing, on the Second Lancet Study regarding Iraq Casualties. <http://musingsoniraq.blogspot.com/2012/08/the-major-flaws-with-lancet-reports-on.html>
432. M.J. van der Laan (2008), "The Construction and Analysis of Adaptive Group Sequential Designs" (March 2008). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 232. <http://www.bepress.com/ucbbiostat/paper232>.
433. Mark J. van der Laan, Sherri Rose, and Susan Gruber, "Readings in Targeted Maximum Likelihood Estimation" (September 2009). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 254. <http://www.bepress.com/ucbbiostat/paper254>
434. D. Rubin and M.J. van der Laan (2007), Extending Marginal Structural Models through Local, Penalized, and Additive Learning (September 2006). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 212. <http://www.bepress.com/ucbbiostat/paper212> Local marginal structural models.
435. Merrill D. Birkner, Alan E. Hubbard, and Mark J. van der Laan, "Data Adaptive Pathway Testing" (November 2005). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 197. <http://www.bepress.com/ucbbiostat/paper197> .
436. M.J. van der Laan (1994), Proving Efficiency of NPMLE and Important Identities. Working Paper #44, Division of Biostatistics, UC Berkeley.
437. A. Molinaro, M. J. van der Laan, D. H. Moore, and K. Kerlikowske (2004), Survival Point Estimate Prediction in Cohorts with Nested Case-Control Study Designs.
438. M.J. van der Laan (1997), Locally Efficient Estimation with Current Status Data. *1996 Proceedings, Biometrics Section of the Annual Meeting of the American Statistical Association, Chicago*, 41–49.
439. M.J. van der Laan and D.R. Peterson (1997), Smooth Estimation and Inference with Interval Censored Data, Working Paper #66, Division of Biostatistics, UC Berkeley.
440. D.R. Peterson, M.J. van der Laan (2002), Local Polynomial Density Estimation with Interval Censored Data. Technical report, Division of Biostatistics, UC Berkeley.
441. T.A. Henneman, M.J. van der Laan, A.E. Hubbard (2002), Estimating Causal Parameters in Marginal Structural Models with Unmeasured Confounders Using Instrumental Variables, Technical report, U.C. Berkeley Division of Biostatistics Working Paper Series. Working paper 104.
442. M.J. van der Laan (1998), Nonparametric Maximum Likelihood. *Encyclopedia of Biostatistics, Survival Analysis*, **11**, John Wiley & Sons, 1998.
443. M.J. van der Laan (1998), Estimation with Interval Censored Data in Longitudinal Studies. Working Paper #74, Division of Biostatistics, UC Berkeley.

444. M.J. van der Laan, R.D. Gill, and J.M. Robins (2000), Locally efficient estimation in censored data models, Theory and Examples, Working Paper #85, Division of Biostatistics, UC Berkeley.
445. K. Pollard, M.J. van der Laan (2002), A method to identify significant clusters in gene expression data, Invited Proceedings of Sci2002, July 2002, Volume II, 318–325.
446. A. Molinaro, M.J. van der Laan, D. Moore (2002), Comparative genomic hybridization array analysis. Working Paper #106, Division of Biostatistics, University of California, Berkeley.
447. K. Pollard, M.J. van der Laan (2002), New methods for identifying significant clusters in gene expression data, Proceedings of JSM 2002.
448. Z. Yu, M.J. van der Laan (2003), Double robust estimation in longitudinal marginal structured models. Working Paper #132, Division of Biostatistics, UC Berkeley.
449. S. Dudoit and M. J. van der Laan (2003), Unified cross-validation methodology for estimator selection and applications to genomics. Bulletin of the International Statistical Institute, 54th Session Proceedings, Vol. LX, Book 2, p. 412-415.
450. van der Laan, M.J., Dudoit (2003), Unified Cross-Validation Methodology For Selection among Estimators, and a General Cross-validated Adaptive epsilon-Net Estimator: Finite Sample Oracle Inequalities and Examples, Working Paper #130, Division of Biostatistics, UC Berkeley.
451. K.S. Pollard, M.J. van der Laan (2003), Multiple testing for gene expression data: an investigation of null distributions with consequences for the permutation test, *Proceedings, 2003 International MultiConference in Computer Science and Engineering, METMBS'03 Conference*, pp. 3-9.
452. Z. Yu and M.J. van der Laan (2003), Measuring Treatment Effects Using Semiparametric Models, Working Paper #136, Division of Biostatistics, UC Berkeley.
453. T. Haight, R. Neugebauer, I. B. Tager, and M. J. van der Laan (2003), Comparison of the Inverse Probability of Treatment Weighted (IPTW) Estimator With a Naive Estimator in the Analysis of Longitudinal Data With Time-Dependent Confounding: A Simulation Study, Working Paper #140, Division of Biostatistics, UC Berkeley.
454. A. Chambaz, W. Zheng, and M.J. van der Laan. Data-adaptive inference of the optimal treatment rule and its mean reward. The masked bandit. Technical report, U.C. Berkeley Division of Biostatistics Working Paper Series, <https://biostats.bepress.com/ucbbiostat/paper349/>, 2016.
455. † C. Ju, D. Benkeser, and M. J. van der Laan. Flexible Collaborative Estimation of the Average Causal Effect of a Treatment using the Outcome-Highly-Adaptive Lasso. ArXiv e-prints, June 2018.
456. . Ju, A. Chambaz, and M.J. van der Laan. Collaborative targeted inference from continuously indexed nuisance parameter estimators. ArXiv e-prints, March 2018.
457. C. Ju, J. Schwab, and M.J. van der Laan. On Adaptive Propensity Score Truncation in Causal Inference. ArXiv e-prints, July 2017.
458. E. Ogburn, O. Sofrygin, I. Diaz, and M. van der Laan. Causal inference for social network data. ArXiv e-prints, arXiv:1705.08527v4. May 2017.
459. O. Sofrygin, Z. Zhu, J. A Schmittdiel, A. S. Adams, R. W. Grant, M. J. van der Laan, and R. Neugebauer. Targeted Learning with Daily EHR Data. ArXiv e-prints, May 2017.

460. M. van der Laan and A. Bibaut. Uniform Consistency of the Highly Adaptive Lasso Estimator of Infinite Dimensional Parameters. ArXiv e-prints, September 2017.
461. W. Zheng, A. Chambaz, and M. J. van der Laan. Drawing valid targeted inference when covariate-adjusted response-adaptive rct meets data-adaptive loss-based estimation, with an application to the LASSO. Technical Report 339, U.C. Berkeley Division of Biostatistics Working Paper Series, 2015.
462. M.J. van der Laan (2017), Finite sample inference for targeted learning, arXiv:1708.09502
463. Tyler J. VanderWeele, Alex R. Luedtke, Mark J. van der Laan, Ronald C. Kessler. Selecting optimal subgroups for treatment using many covariates. arXiv:1802.09642.
464. Caleb H. Miles, Maya Petersen, Mark J. van der Laan Causal Inference for a Single Group of Causally-Connected Units Under Stratified Interference. arXiv:1710.09588.
465. Weixin Cai, Mark J. van der Laan (2018), One-step Targeted Maximum Likelihood for Time-to-event Outcomes. arXiv:1802.09479
466. Oleg Sofrygin, Zheng Zhu, Julie A Schmittiel, Alyce S. Adams, Richard W. Grant, Mark J. van der Laan, Romain Neugebauer Targeted Learning with Daily EHR Data. arXiv:1705.09874.
467. Kara E. Rudolph, Oleg Sofrygin, Wenjing Zheng, Mark J. van der Laan (2018), Robust and flexible estimation of data-dependent stochastic mediation effects: a proposed method and example in a randomized trial setting. arXiv:1707.09021.
468. Aurelien F. Bibaut, Mark J. van der Laan Data-adaptive smoothing for optimal-rate estimation of possibly non-regular parameters. arXiv:1706.07408.
469. Ju, Cheng, Mary Combs, Samuel D. Lendle, Jessica M. Franklin, Richard Wyss, Sebastian Schneeweiss, and Mark J. van der Laan. "Propensity score prediction for electronic healthcare databases using Super Learner and High-dimensional Propensity Score Methods." arXiv preprint arXiv:1703.02236 (2017).
arXiv:1703.02237