

EXPERIENCE

Co-Founder, TL Revolution LLC 2022–present

Develop software to implement Targeted Learning methodologies in a regulatory setting, with emphasis on producing robust real-world evidence in support of decision making regarding safety and efficacy. We also offer training courses and strategic consulting partnerships.

Principal, Putnam Data Sciences 2017–present

Provide statistical consulting and training in methodology for statistical causal inference and predictive modeling. Clients and collaborators include the U.S. Food and Drug Administration, Harvard University, Brown University, Stanford University, University of California at Berkeley, University of Utah, and Kaiser Permanente Division of Research.

Director of Biostatistics Center and Assistant Professor 2016–2017

Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute

Methodologic research interests include targeted learning for causal inference, methods for detecting safety signals in electronic health data, and applications of machine learning in predictive modeling and propensity score estimation. Applied work included an analysis of sepsis trends in the United States, evaluating vaccine safety with respect to risk of febrile seizure, and studying the effects of short vs. long course antibiotic use in patients with ventilator-acquired pneumonia. Conducted a successful faculty search culminating in a new hire at the Associate Professor level.

Senior Director, IMEDS-Methods Research 2014–2015

Reagan-Udall Foundation for the FDA

Developed and executed a research agenda aimed at improving methods for safety surveillance using large electronic healthcare databases. Focus areas include methods for addressing bias in observational studies, understanding sources of heterogeneity in observational studies using administrative data, and extracting lessons from activities of the Observational Medical Outcomes Partnership (OMOP) relevant to Mini-Sentinel. Oversight of communication and dissemination among a broad set of stakeholders from industry, academia, and the FDA.

Post-Doctoral Research Fellow, Harvard School of Public Health 2011–2014

Department of Epidemiology, Mentors: Miguel Hernán and James Robins

Investigated theoretical and computational issues surrounding data-adaptive estimation of weights for inverse probability weighting estimators in large longitudinal datasets. Development and comparison of data adaptive methodology for causal inference, including alternative double-robust estimators. Presented a critique of the Observational Medical Outcomes Partnership (OMOP) work on identifying adverse drug events in large observational datasets to Mini-Sentinel leadership at the U.S. Food and Drug Administration.

Technical Consultant , Ordinology Consulting Technical editor for technical books and papers and database applications development for tracking system for foster children's medical records.	1989–2004
Instructor , Computer Information Systems Department, De Anza College <i>Programming in C, Introduction to Computing, Scientific Programming</i>	1992–1997
Research Assistant , Computer Science Department, UC San Diego Database management system interface development	1988–1989
Teaching Assistant , Computer Science Department, UC San Diego <i>Artificial Intelligence I and II, Comparative Study of Programming Languages, Combinatorics and Graph Theory</i>	1988–1989
Instructor , Computer Science Department, UC San Diego <i>Introduction to Programming Techniques</i>	1988
Scientific Programmer , Department of Applied Environmental Sciences, SAIC Designed and implemented a Laboratory Information Management System, developed statistical analysis and plotting tools in SAS	1985–1987

EDUCATION

<i>Ph.D.</i> , Biostatistics, University of California, Berkeley Advisor: Mark J. van der Laan Dissertation: <i>Collaborative Targeted Maximum Likelihood Estimation</i> awarded the <i>Erich L. Lehmann Citation for Outstanding Ph.D. Dissertation in Theoretical Statistics</i>	2011
<i>M.P.H.</i> , Epidemiology and Biostatistics, University of California, Berkeley Master's Paper: <i>Survival Analysis: A Comparison of Clustering and Parametric Approaches</i>	2007
<i>M.S.</i> , Computer Science, University of California, San Diego Master's Paper: <i>A Conceptual Framework for Temporal Databases</i>	1989
<i>B.A.</i> , Chemistry, Northwestern University, Evanston, Illinois	1982

PUBLICATIONS

Peer Reviewed

1. **S Gruber**, RV Phillips, H Lee, M Ho, J Concato, MJ van der Laan. Targeted learning: Toward a future informed by real-world evidence. *Statistics in Biopharmaceutical Research*; DOI: 10.1080/19466315.2023.2182356.
2. RV Phillips, MJ van der Laan, H Lee, **S Gruber**. Practical considerations for specifying a super learner. *International Journal of Epidemiology*, 2023; dyad023.

3. M Ho, **S Gruber**, Y Fang, DE Faris, P Mishra-Kalyani, D Benkeser M van der Laan. Examples of Applying RWE Causal-Inference Roadmap to Clinical Studies. *Statistics in Biopharmaceutical Research* 2023; DOI: 10.1080/19466315.2023.2177333.
4. DS Carrell, **S Gruber**, JS Floyd, MA Bann, KL Cushing-Haugen, RL Johnson, V Graham, DJ Cronkite, BL Hazlehurst, AH Felcher, CA Bejan, A Kennedy, M Shinde, S Karami, Y Ma, D Stojanovic, Y Zhao, R Ball, J Nelson, Improving Methods of Identifying Anaphylaxis for Medical Product Safety Surveillance Using Natural Language Processing and Machine Learning. *American Journal of Epidemiology*, 2022; kwac182, <https://doi.org/10.1093/aje/kwac182>.
5. CW Kahler, T Janssen, **S Gruber**, CJ Howe, MB Laws, J Walthers, M Magill, NR Mastroleo, PM Monti. Change talk subtypes as predictors of alcohol use following brief motivational intervention. *Psychology of Addictive Behaviors*, 2022. <https://doi.org/10.1037/adb0000898>.
6. **S Gruber**, H Lee, R Phillips, M Ho, M van der Laan. Developing a Targeted Learning-Based Statistical Analysis Plan. *Statistics in Biopharmaceutical Research* 2022; <https://doi.org/10.1080/19466315.2022.2116104>.
7. **S Gruber**, RV Phillips, H Lee, MJ van der Laan. Data-Adaptive Selection of the Propensity Score Truncation Level for Inverse Probability Weighted and Targeted Maximum Likelihood Estimators of Marginal Point Treatment Effects. *American Journal of Epidemiology* 2022; 191(9), 1640–1651, <https://doi.org/10.1093/aje/kwac087>.
8. **S Gruber**, DS Carrell, JS Floyd, JC Nelson, BL Hazlehurst, PJ Heagerty. Letter to the Editor re Beachler, et al. *Pharmacoepidemiology and Drug Safety* 2021; 30 (12), 1735-1736.
9. MA Bann, DS Carrell, **S Gruber**, M Shinde, R Ball, JC Nelson, JS Floyd. Identification and validation of anaphylaxis using electronic health data in a population-based setting. *Epidemiology* 2021; 32 (3), 439-443.
10. HR Elder, **S Gruber**, SJ Willis, N Cocoros, M Callahan, EW Flagg, M Klompas, KK Hsu. Can machine learning help identify patients at risk for recurrent sexually transmitted infections? *Sexually Transmitted Diseases* 2021; 48 (1), 56-62.
11. B Fireman, **S Gruber**, Z Zhang, R Wellman, JC Nelson, J Franklin, J Maro, C Rogers Murray, S Toh, J Gagne, S Schneeweiss, L Amsden, R Wyss. Consequences of depletion of susceptibles for hazard ratio estimators based on propensity scores. *Epidemiology* 2020; 31(6).
12. **S Gruber**, D Krakower, JT Menchaca, K Hsu, R Hawrusik, JC Maro, NM Cocoros, BA Kruskal, IB Wilson, KH Mayer, M Klompas. Using electronic health records to identify candidates for human immunodeficiency virus pre-exposure prophylaxis: An application of super learning to risk prediction when the outcome is rare. *Statistics in Medicine* 2020; 39 (23), 3059-3073.
13. MA Baker, C Jankosky, WK Yih, **S Gruber**, L Li, NM Cocoros, H Lipowicz, C Coronel-Moreno, S DeLuccia, ND Lin, CN McMahill-Walraven, D Menschik, MS Selvan, N Selvam, RC Tilney, L Zichittella, GM Lee, AT Kawai. The risk of febrile seizures following influenza and 13-valent pneumococcal conjugate vaccines. *Vaccine* 2020; 38 (9), 2166-2171.
14. DS Krakower, **S Gruber**, K Hsu, JT Menchaca, JC Maro, BA Kruskal, IB Wilson, KH Mayer, M Klompas. Development and validation of an automated HIV prediction algorithm to identify candidates for pre-exposure prophylaxis: a modelling study. *Lancet HIV* 2019; 6 (10), e696-e704.
15. **S Gruber**, MJ van der Laan. Comment on Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition. *Statistical Science* 2019; 34(1): 82-85. DOI: 10.1214/18-STS689.

16. A Weber, GL Darmstadt, **S Gruber**, ME Foeller, SL Carmichael, DK Stevenson, GM Shaw. Application of machine-learning to predict early spontaneous preterm birth among nulliparous non-Hispanic black and white women. *Annals of Epidemiology* 2018; 28(11): 783-789.e1; <https://doi.org/10.1016/j.annepidem.2018.08.008>.
17. K Yoshida, **S Gruber**, BH Fireman, S Toh. Comparison of privacy-protecting analytic and data-sharing methods: A simulation study. *Pharmacoepidemiology and Drug Safety* 2018;27 (9):1034-41.
18. D Martin, JJ Gagne, **S Gruber**, RIzem, JC Nelson, MD Nguyen, R Ouellet Hellstrom, S Schneeweiss, S Toh, AM Walker. Sequential surveillance for drug safety in a regulatory environment. *Pharmacoepidemiology and Drug Safety* 2018.
19. SV Wang, Y Jin, B Fireman, **S Gruber**, M He, R Wyss, HJ Shin, Y Ma, S Keeton, S Karami, JM Major, S Schneeweiss, JJ Gagne. Relative performance of propensity score matching strategies for subgroup analyses. *Epidemiology* 2018.
20. C Rhee, R Dantes, L Epstein, DJ Murphy, CW Seymour, TJ Iwasyna, SS Kadri, DC Angus, RL Danner, AE Fiore, JA Jernigan, GS Martin, E Septimus, DK Warren, A Karcz, C Chan, JT Menchaca, R Wang, **S Gruber**, M Klompas. Incidence and Trends of Sepsis in US Hospitals Using Clinical vs Claims Data, 2009-2014. *Journal of the American Medical Association* 2017; 318 (13), 1241-1249.
21. C Ju, **S Gruber**, SD Lendle, A Chambaz, JM Franklin, R Weiss, S Schneeweiss, MJ van der Laan. Scalable Collaborative Targeted Learning for High-Dimensional Data. *Statistics in Medicine* 2017; <https://doi.org/10.1177/0962280217729845>.
22. M Klompas, L Li, JT Menchaca, **S Gruber**. Ultra-Short-Course Antibiotics for Patients With Suspected Ventilator-Associated Pneumonia but Minimal and Stable Ventilator Settings. *Clinical Infectious Disease* 2016; 64 (7), 870-876.
23. **S Gruber**, A Chakravarty, SR Heckbert, M Levenson, D Martin, J Nelson, B Psaty, S Pinheiro, C Reich, S Schneeweiss, S Toh, A Walker. Design and Analysis Choices for Safety Surveillance Evaluations Need to be Tuned to the Specifics of the Hypothesized Drug-Outcome Association. *Pharmacoepidemiology and Drug Safety* 2016; 25 (9), 2315-2336.
24. **S Gruber**, E Tchetgen Tchetgen. Limitations of empirical calibration of p-values using observational data. *Statistics in Medicine* 2016; 35 (22), 3869-3882.
25. MJ van der Laan and **S Gruber**. One-Step Targeted Minimum Loss-based Estimation Based on Universal Least Favorable One-Dimensional Submodels. *The International Journal of Biostatistics* 2016; 12 (1), 351 - 378.
26. **S Gruber**. Targeted Learning in Healthcare Research. *Big Data* 2016; 3(4), 211-218. DOI: 10.1089/big.2015.0025.
27. **S Gruber**, RW Logan, I Jarrín, S Monge, MA Hernán. Ensemble learning of inverse probability weights for marginal structural modeling in large observational datasets. *Statistics in Medicine* 2015; 34 (1), 106-117.
28. **S Gruber**. A Causal Perspective on OSIM2 Data Generation, with Implications for Simulation Study Design and Interpretation. *Journal of Causal Inference* 2015. ISSN (Online) 2193-3685, ISSN (Print) 2193-3677, DOI: 10.1515/jci-2014-0008.
29. ME Schnitzer, JJ Lok, **S Gruber**. Variable Selection for Confounder Control, Flexible Modeling and Collaborative Targeted Minimum Loss-Based Estimation in Causal Inference. *The International Journal of Biostatistics* 2015; ISSN (Online) 1557-4679, ISSN (Print) 2194-573X, DOI: 10.1515/ijb-2015-0017.

30. C Bell, A Chakravarty, **S Gruber**, S Heckbert, M Levenson, D Martin, J Nelson, S Pinheiro, B Psaty, C Reich, S Schneeweiss, A Shoaibi, S Toh, A Walker. Characteristics of study design and elements that may contribute to the success of electronic safety monitoring systems. *Pharmacoepidemiology and Drug Safety* 2014; 23 (11).
31. M Petersen, J Schwab, **S Gruber**, N Blaser, M Schomaker, M van der Laan. Targeted maximum likelihood estimation for dynamic and static longitudinal marginal structural working models. *Journal of Causal Inference* 2014; 2(2), 147-185.
32. N Krief, **S Gruber**, R Radice, R Grieve, JS Sekhon. Evaluating treatment effectiveness under model misspecification: a comparison of targeted maximum likelihood estimation with bias-corrected matching. *Statistical Methods in Medical Research* (2014) [Epub ahead of print].
33. **S Gruber** and MJ van der Laan. An Application of Targeted Maximum Likelihood Estimation to the Meta-Analysis of Safety Data. *Biometrics* 2013; [doi: 10.1111/j.1541-0420.2012.01829.x].
34. PB Ryan, MJ Schuemie, **S Gruber**, I Zorych, D Madigan. Empirical performance of the new user cohort design: Lessons for developing a risk identification and analysis system. *Drug Safety* (2013) 36 (Suppl 1):S59DS72.
35. **S Gruber** and MJ van der Laan. tmle: An R Package for Targeted Maximum Likelihood Estimation. *Journal of Statistical Software* 2012; 51(13).
36. **S Gruber** and MJ van der Laan. Targeted Minimum Loss Based Estimation of a Causal Effect on an Outcome with Known Conditional Bounds. *The International Journal of Biostatistics* 2012; 8(1).
37. MJ van der Laan and **S Gruber**. Targeted Minimum Loss Based Estimation of Causal Effects of Multiple Time Point Interventions. *The International Journal of Biostatistics* 2012; 8(1).
38. **S Gruber** and MJ van der Laan. Targeted Minimum Loss Based Estimator that Outperforms a Given Estimator. *The International Journal of Biostatistics* 2012; 8(1), Article 11.
39. KE Porter*, **S Gruber***, MJ van der Laan, and JS Sekhon. The Relative Performance of Targeted Maximum Likelihood Estimators. *The International Journal of Biostatistics* 2011; 7(1), Article 31.
40. ML Petersen, KE Porter, **S Gruber**, Y Wang, MJ van der Laan. Diagnosing and Responding to Violations of the Positivity Assumption. *Statistical Methods in Medical Research*, 2010.
41. **S Gruber** and MJ van der Laan. A Targeted Maximum Likelihood Estimator of a Causal Effect on a Bounded Continuous Outcome. *The International Journal of Biostatistics* 2010; 6(1), Article 26.
42. MJ van der Laan and **S Gruber**. Collaborative Double Robust Targeted Maximum Likelihood Estimation. *The International Journal of Biostatistics* 2010; 6(1), Article 17.
43. **S Gruber** and MJ van der Laan. An Application of Collaborative Targeted Maximum Likelihood Estimation in Causal Inference and Genomics. *The International Journal of Biostatistics* 2010; 6(1), Article 18.
44. NL Fleischer, AM Weber, **S Gruber**, KZ Arambula, M Mascarenhas, JA Frasure, C Wang, SL Syme. Pathways to Health: A Framework for Health-Focused Research and Practice. *Emerging Themes in Epidemiology* 2006; 3:18.

Book Chapters

1. **S Gruber**, H Lee, RV Phillips, and MJ van der Laan. Causal Inference with Targeted Learning for Producing and Evaluating Real-World Evidence. In W He, Y Fang and H Wang ed, *Real-World Evidence in Medical Product Development. 2023, Springer, New York.*
2. **S Gruber** and MJ van der Laan. Collaborative Targeted Maximum Likelihood Estimation to assess causal effects in observational studies . In KE Peace ed, *Biopharmaceutical Applied Statistics Symposium (BASS), Volume 2, Chapter 2. 2018, Springer, New York.*
3. **S Gruber**. Targeted Learning with Application to Health Care Research. In E Sejdic and TH Falk eds, *Biomedical Signal Processing in Big Data, 28, CRC Press, Florida.*
4. **S Gruber** and MJ van der Laan. A Targeted Maximum Likelihood Estimator of a Causal Effect on a Bounded Continuous Outcome. In MJ van der Laan and S Rose, *Targeted Learning: Prediction and Causal Inference for Observational and Experimental Data, Chapter 7. 2011, Springer, New York.*
5. ML Petersen, KE Porter, **S Gruber**, Y Wang, and MJ van der Laan. Diagnosing and Responding to Violations in the Positivity Assumption. In MJ van der Laan and S Rose, *Targeted Learning: Prediction and Causal Inference for Observational and Experimental Data, Chapter 10. 2011, Springer, New York.*
6. **S Gruber** and MJ van der Laan. Collaborative Targeted Maximum Likelihood Estimation. In MJ van der Laan and S Rose, *Targeted Learning: Prediction and Causal Inference for Observational and Experimental Data, Chapter 19. 2011, Springer, New York.*
7. JS Sekhon, **S Gruber**, KE Porter and MJ van der Laan. Doubly Robust Estimators and the Neyman-Rubin Causal Model. In MJ van der Laan and S Rose, *Targeted Learning: Prediction and Causal Inference for Observational and Experimental Data, Chapter 21. 2011, Springer, New York.*

Non Peer-Reviewed

1. LE Dang, **S Gruber**, et al. A Causal Roadmap for Generating High-Quality Real-World Evidence 2023; arXiv preprint arXiv:2305.06850.
2. BD Williamson, et al. An Application of the Causal Roadmap in Two Safety Monitoring Case Studies: Covariate-Adjustment and Outcome Prediction using Electronic Health Record Data. 2023; arXiv preprint arXiv:2305.07564.
3. **S Gruber**, RV Phillips, H Lee, J Concato, MJ van der Laan. Evaluating and improving real-world evidence with Targeted Learning 2022; arXiv preprint arXiv:2208.07283.
4. JS Floyd, **S Gruber**, DS Carrell, MA Bann. Re. Identification and Validation of Anaphylaxis Using Electronic Health Data in a Population-based Setting Respond. *Epidemiology* 2022; 33(1), E2-E3.
5. **S Gruber**, B Fireman, R Izem. Performance of Different Propensity Score Methods in Simulated Cohort Studies with Time-to-Event Outcomes, 2019. Published online available at: <https://www.sentinelinitiative.org>.
6. M Klompas and **S Gruber**. Reply to Boyer et al. *Clinical Infectious Diseases*. 2017 Jun 15;64(12):1803-4.
7. MA Baker, C Jankosky, K Yih, **S Gruber**, L Li, NM Cocoros, H Lipowicz, C Coronel-Moreno, S Feibelmann, N Lin, CN McMahill-Walraven, D Menschik, MS Selvan, N Selvam, RC Tilney,

- L Zichittella, GM Lee, A T Karai. Influenza Vaccines and Febrile Seizures in the 2013-2014 and 2014-2015 Influenza Seasons. Sentinel CBER/PRISM Surveillance Report, 2017.
8. **S Gruber** and MJ van der Laan. Consistent Causal Effect Estimation Under Dual Misspecification and Implications for Confounder Selection Procedures. *Statistical Methods in Medical Research* [epub ahead of print February, 2012].
 9. **S Gruber** and MJ van der Laan. Targeted Maximum Likelihood Estimation: A Gentle Introduction. UC Berkeley Division of Biostatistics Working Paper Series. Working Paper 252 (2009).

PRESENTATIONS, WEBINARS, WORKSHOPS, AND SHORT COURSES

Invited

1. Improving medical product safety surveillance with natural language processing and machine learning. *Regulatory Challenges, Prospects, and Modern Solutions*. Regulatory Affairs Professional Society. June, 2023.
2. *Understanding the Use of Negative Controls to Assess the Validity of Non-Interventional Studies of Treatment Using Real-World Evidence*. Panelist. Duke-Margolis Center for Public Health Policy. March, 2023.
3. A Targeted Learning Framework for Causal Effect Estimation Using Real-World Data. *78th Deming Conference*, Philadelphia, PA, December, 2022.
4. Enterprise software for learning from real-world data: There's no time like the present. *Forum on the Integration of Observational and Randomized Data*, Washington D.C. November, 2022.
5. Short course: A Targeted Learning Framework for Causal Effect Estimation using Real-World Data. U.S. Food and Drug Administration, Silver Spring, MD, March, 14-18, 2022.
6. Electronic Health Records + Natural Language Processing + Machine Learning = Improved Sentinel Outcome Detection Algorithms. *Great Lakes Pharmacovigilance and Drug Safety Forum, University of Illinois*, Chicago, IL, October, 2021.
7. Short course: Beyond Propensity Score Methods for Real-World Evidence: Targeted Maximum Likelihood Estimation (TMLE) with an Ensemble of Machine Learning Algorithms. *2021 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop*, September 21, 2021.
8. Developing a Targeted Learning-Based Statistical Analysis Plan. *Targeted Learning Webinar Series*, Putnam Data Sciences. April 29, 2021.
9. Electronic Health Records + Natural Language Processing + Machine Learning = Improved Sentinel Outcome Detection Algorithms. *US Food and Drug Administration Sentinel Innovation and Methods Seminar Series*, Silver Spring, MD, Nov 19, 2020.
10. An Introduction to Targeted Maximum Likelihood Estimation of Causal Effects. Targeted Learning Webinar Series, Putnam Data Sciences. March 3, 2020.
11. Targeted Learning for Estimating Causal Effects in Observational and Randomized Studies. *ISPE Webinar*, January 2020.
12. Machine Learning in Causal Inference: The Role of domain knowledge. *Dean's Symposium, Statistics and the Life Sciences: Creating a Healthier World*, Boston University School of Public Health and the Department of Mathematics and Statistics, Boston, MA, November 15, 2019.

13. A Targeted Learning Perspective on Pharmacovigilance. *Causal Inference Seminar, Department of Biostatistics and Statistics*, Boston University, Boston, MA, April 22, 2019.
14. Causal Thinking in post-market assessments of drug safety. *Boston University Student Chapter of the ASA Statistics in Practice Panel*, April 17, 2019, Boston, MA.
15. Targeted Learning for Causal Inference. *Department of Biostatistics, Harvard TH Chan School of Public Health*, Boston, Massachusetts, November, 2018.
16. Targeted Learning for Data Adaptive Causal Inference in Observational and Randomized Studies. *Third Seattle Symposium on Healthcare Data Analytics*, Seattle, Washington, October, 2018.
17. Beyond Logistic Regression: Machine Learning for Propensity Score Estimation. *FDA CDER Office of Biostatistics Division of Biometrics VII*, Silver Spring, Maryland, January, 2018.
18. A Discussion of P-Value Calibration. *FDA Office of Surveillance and Epidemiology Sentinel Webinar*, Silver Spring, Maryland, January, 2018.
19. Performance of Different Propensity Score Methods in Simulated Cohort Studies with Time-to-Event Outcomes. *The 2018 International Conference on Health Policy Statistics*, Charleston, SC, January, 2018.
20. An Introduction to Super Learning for Prediction. *Takeda Pharmaceuticals, Inc.*, Boston, Massachusetts, June, 2017.
21. Using Electronic Health Records to Identify Candidates for HIV Pre-Exposure Prophylaxis: An Application of Super Learning to Risk Prediction when the Outcome is Rare. *Q-GRID Working Group Seminar Series*, Harvard TH Chan School of Public Health, Boston, Massachusetts, January, 2017.
22. Targeted Learning for data adaptive causal inference in observational and randomized studies. *2nd Seattle Symposium on Healthcare Data Analytics*, Seattle, WA, October, 2016.
23. Targeted Learning in Health Care Research. *Harvard T.H. Chan School-Takeda Symposium*, Boston, Massachusetts, September, 2016.
24. Considerations for Covariate Selection When Using TMLE. *Atlantic Causal Inference Conference*, New York, New York, May, 2016.
25. Keynote Panelist, Computer Power and Human Reason: From Calculation to Judgment. *31st International Conference on Pharmacoepidemiology and Therapeutic Risk Management*, Boston, Massachusetts, August, 2015.
26. Insights into the OMOP 2011-2012 Experiment and P-value Calibration. *FDA Office of Surveillance and Epidemiology Safety Seminar*, November, 2015.
27. An Overview of Targeted Maximum Likelihood Estimation. *Group Health Research Institute*, November, 2015.
28. An Introduction to Longitudinal TMLE. *University of Utah School of Medicine*, November, 2015.
29. IMEDS-Evaluation for Conducting Post-Market Safety Studies. *Big DiP USA 2015*, Boston, Massachusetts, September, 2015.
30. Computation and Big Healthcare Data (keynote panelist). *31st International Conference on Pharmacoepidemiology and Therapeutic Risk Management*, Boston, Massachusetts, August, 2015.
31. Big Data and Safety Surveillance: Are We Building a Bigger Haystack? *Drug Information Association (DIA) 51st Annual Meeting*, Washington, D.C., June, 2015.

32. Can P-Value Calibration Really Produce More Reliable Evidence for Safety Surveillance? *IMEDS Community Call Webinar*, November, 2015.
33. Recent Work of the Joint Mini-Sentinel/IMEDS Work Group. *Mini-Sentinel Investigator Meeting, US Food and Drug Administration*, November, 2014.
34. Gaps and Opportunities: Methodologic Challenges in Post-Market Safety Surveillance. *First Seattle Symposium on Healthcare Data Analytics*, Seattle, Washington, September, 2014.
35. Methods to Adjust for Treatment Switching that Disrupts Randomization in Clinical Trials. *Joint Statistical Meetings*, Boston, Massachusetts, August, 2014.
36. A Targeted Learning Approach to Mitigating Bias in Effect Estimates when Some Outcome Data are Missing. *SER Symposium*, Seattle, Washington, June, 2014.
37. Active Prescription Drug Safety Surveillance: Exploring OMOP 2011-2012 Experiments. *GNS Health Care*, Cambridge, Massachusetts, October, 2013.
38. Active Prescription Drug Safety Surveillance: Exploring OMOP 2011-2012 Experiments. *Division of Pharmacoepidemiology and Pharmacoeconomics, Harvard Medical School*, September, 2013.
39. Introduction to Targeted Learning. *Spiegelman Group, Department of Epidemiology, Harvard School of Public Health*, August, September, November, 2013. Boston, Massachusetts.
40. Collaborative Triple-Robust Targeted Loss-Based Estimation. *The Western North American Region of the International Biometric Society*, June, 2013. Los Angeles, California.
41. Efficient Semiparametric Modeling in Causal Inference. *Annual Meeting of the Statistical Society of Canada*, May, 2013. Edmonton, Canada.
42. Workshop on Targeted Maximum Likelihood Estimation. *Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill*, February, 2013. Chapel Hill, North Carolina.
43. Methodological Advances in Causal Inference: Addressing Challenges in Comparative Effectiveness Research and Identification of Adverse Drug Effects. **S Gruber**. *Division of Pharmacoepidemiology and Pharmacoeconomics, Harvard Medical School*, October, 2012.
44. Developments in Causal Inference, What's New and What's Relevant for Health Economic Evaluation? R Grieve, R Radice, Z Sadique, N Krief, R Ramsahai, **S Gruber**. *Methods for Addressing Selection Bias in Health Economic Evaluation, The London School of Hygiene and Tropical Medicine*, (Short Course) July, 2012. London, England.
45. An Application of Targeted Maximum Likelihood Estimation to Health Economic Evaluation. **S Gruber**. *Centre for Statistical Methodology, The London School of Hygiene and Tropical Medicine*, July, 2012. London, England.
46. Collaborative Targeted Maximum Likelihood Estimation. **S Gruber** and MJ van der Laan. *State of the Art Methodology in Causal Analysis, WNAR 2011 IMS Annual Meeting*, June 20, 2011. San Luis Obispo, CA.
47. Targeted Maximum Likelihood Estimation for Estimating Causal Effects. **S Gruber** and MJ van der Laan. *Quantitative Sciences Research Seminar, Department of Medicine, Stanford University*, June 7, 2011. Palo Alto, CA.
48. Implications of the Collaborative Double Robustness Property of the Efficient Influence Curve for Nuisance Parameter Estimation. **S Gruber** and MJ van der Laan. *Causal Inference Group Meeting, Harvard School of Public Health*. January 19, 2011. Boston, MA.

49. Application of Targeted Maximum Likelihood Estimation to the Meta-Analysis of Randomized Clinical Trials Data, Safety and Efficacy. MJ van der Laan, R Neugebauer, **S Gruber** *Causal Inference Workshop, US Food and Drug Administration*, September, 2009. Silver Spring, MD.

Contributed

50. Automated Identification of Potential Candidates for HIV Pre-Exposure Prophylaxis using Electronic Health Record Data. *Brown Bag Series, Department of Population Medicine*, Boston, Massachusetts, March, 2017.
51. Extracting information from Observational Electronic Health and Claims Data to Enhance Post-Approval Medical Product Safety Surveillance. **S Gruber** and P Ryan. Roundtable Discussion, FDA-Industry Workshop, September, 2014. Washington, DC.
52. Ensemble learning of Inverse Probability Weights for Marginal Structural Modeling in Large Observational Datasets. **S Gruber**, RW Logan, I Jarrín, S Monge, MA Hernán *Eastern North American Region/International Biometric Society Spring Meeting (ENAR)*, March, 2014. Baltimore, MD.
53. Evaluating Treatment Effectiveness Under Model Misspecification: A Comparison of Targeted Maximum Likelihood Estimation with Bias-Corrected Matching. N Kreif, **S Gruber**, R Radice, R Grieve, and JS Sekhon. *Joint Statistical Meetings*, August, 2013. Montreal, Canada.
54. Investigating the Impact of OSIM2 Data Generation on Estimator Performance. **S Gruber**. *OMOP Stats working group meeting*, July, 2013. Webinar.
55. Targeted Maximum Likelihood Estimation with Known Conditional Bounds. **S Gruber** and MJ van der Laan. *Eastern North American Region/International Biometric Society Spring Meeting (ENAR)*, March, 2013. Orlando, FL.
56. Effect of propensity score adjustment on risk ratio estimates in OMOP cohort method analyses. **S Gruber** and J Robins. *OMOP Stats working group meeting*, October, 2012. Webinar.
57. Challenges in Data-Adaptive Estimation of Inverse Probability Weights for Causal Effect Estimation in Large Longitudinal Datasets. **S Gruber** and M Hernán. *Joint Statistical Meetings*, July, 2012. San Diego, CA.
58. Targeted Minimum Loss Based Estimation of Causal Effects of Multiple Timepoint Interventions. MJ van der Laan and **S Gruber**. *Eastern North American Region/International Biometric Society Spring Meeting (ENAR)*, April, 2012. Washington, DC.
59. Collaborative Targeted Maximum Likelihood Estimation of Causal Effects. **S Gruber** and MJ van der Laan *Statistical Methods in Epidemiology Working Group, Harvard School of Public Health*, November, 2011. Boston, MA.
60. Software for Targeted Maximum Likelihood Estimation. **S Gruber** and MJ van der Laan. *Joint Statistical Meetings*, August, 2010. Vancouver, BC.
61. The Relative Performance of Collaborative Double Robust Targeted Maximum Likelihood Estimators. KE Porter, **S Gruber**, JS Sekhon, MJ van der Laan. *Joint Statistical Meetings*, August, 2010. Vancouver, BC.
62. Collaborative Targeted Maximum Likelihood Estimation of Causal Effect Parameters. **S Gruber** and MJ van der Laan. *Joint Statistical Meetings*, August, 2009. Washington, DC.

Poster Sessions

63. Investigating the Impact of OSIM2 Data Generation on Estimator Performance. **S Gruber**. *OMOP-IMEDS 2013 Symposium*, November, 2013. Bethesda, MD.

64. A Demonstration of Collaborative Targeted Maximum Likelihood Estimation. **S Gruber** and MJ van der Laan. *Atlantic Causal Inference Conference*, May, 2010. New York, NY.
65. The Relative Performance of Targeted Maximum Likelihood Estimators. KE Porter, **S Gruber**, JS Sekhon, MJ van der Laan. *Atlantic Causal Inference Conference*, May, 2010. New York, NY.
66. Collaborative Targeted Maximum Likelihood Estimation: An Application to Biomarker Discovery. **S Gruber** and MJ van der Laan. *Biomedical Computing at Stanford*, November, 2009. Stanford, CA.

STATISTICAL SOFTWARE

Design, implementation, documentation of statistical software for causal inference, developed in collaboration with Mark van der Laan

tmle_ISM: Software for targeted maximum likelihood estimation of causal effects of multiple time point interventions.

tmle: R package for targeted maximum likelihood estimation of a binary point treatment effect.

tmleLite: R package for targeted maximum likelihood estimation of an additive effect of a binary point treatment. Provides data-adaptive estimation using the *D/S/A* algorithm, influence curve-based inference.

ctmle: Software for collaborative targeted maximum likelihood estimation of an additive point treatment effect, incorporating collaborative data-adaptive nuisance parameter estimation, influence curve-based inference.

bias.pboot: Parametric bootstrap for diagnosing bias due to violations of the positivity assumption (with KE Porter, Y Wang, and M Petersen).

GRANTS, HONORS, AND AWARDS

A Targeted Learning Framework for Causal Effect Estimation Using Real World Data, PI: Susan Gruber. U.S. Food and Drug Administration, Contract 75F40119C10155, Funding Amount: \$411,646. September 28, 2019 - March 28, 2022.

U.S. Patent 8,996,445, *Collaborative targeted maximum likelihood learning*. March 31, 2015

Visiting Scientist, Department of Epidemiology
Harvard T.H. Chan School of Public Health 2014 - 2016

OMOP-IMEDS 2013 Symposium Best Poster Award 2013
Investigating the Impact of OSIM2 Data Generation on Estimator Performance

Association for Women in Mathematics-National Science Foundation Travel Award 2013

Visiting Scientist, Department of Health Services Research and Policy 2012
The London School of Hygiene and Tropical Medicine

Harvard School of Public Health Postdoctoral Association Travel Award 2012

Erich L. Lehmann Citation for an Outstanding Ph.D. Dissertation in Theoretical Statistics (*interpreted broadly*), Department of Statistics, UC Berkeley 2011

The citation reads:

Susan Gruber is honored with this award because of her stellar contributions to the development of collaborative targeted maximum likelihood estimation for causal effect estimation, and integration of software development into the research to push and refine the methods to their fullest potential. Susan Gruber's thesis research also involved applications of the methods to HIV resistance, HIV-clinical trials, and safety analysis in collaboration with the FDA.

Russell M. Grossman Endowment Award, UC Berkeley 2011

France-Berkeley Fund Grant, Chambaz/van der Laan 2010-2011

Bears Breaking Boundaries Award for innovative proposal on neglected diseases and pathways to health (with N Fleischer, A Weber, K Arambula, M Mascarenhas, J Frasure, C Wang) 2006

PROFESSIONAL ACTIVITIES

American Statistical Association Real-World Evidence Biopharmaceutical Working Group member, 2020-present.

Webinar Series developer, producer and host, Targeted Learning Webinar series, Putnam Data Sciences channel, <https://www.youtube.com/channel/UC6Cg1XjzX-MlyxKIWfHezFQ>, 2020-2021.

Data Safety Monitoring Board member, *An adaptive randomized evaluation of nurse-led HIV treatment retention interventions for women living with HIV*, NIH funded sequential multiple assignment randomized trial (SMART), multi-centered. Co-investigators H Hausler (Site PI), S Schwartz, R Deliwe Phetlhu, S Mishra, J Farley, C Beyrer, D Dowdy, E Geng. 2018-2021.

Editorial boards (past service): *The International Journal of Biostatistics*, *Journal of Causal Inference*

Chair, Atlantic Causal Inference Conference Data Challenge Committee, 2019, Montreal, Canada.

Organizing committee member, *Atlantic Causal Inference Conference*, 2013, Cambridge, MA.

Session chair and organizer

- A Targeted Learning Framework for Causal Effect Estimation Using Real-World Data. *DIA 2022 Global Annual Meeting*, June, 2022, Chicago, IL.
- Drug safety surveillance: What's practical now, and where are we heading? *Second Seattle Symposium on Healthcare Data Analytics*, 2016, Seattle, WA.
- Perspectives from Causal Inference in the Era of Big (Healthcare) Data *32nd International Conference on Pharmacoepidemiology and Therapeutic Risk Management*, 2016, Dublin, Ireland.
- Novel Computational Approaches in Safety Surveillance, *Joint Statistical Meetings*, 2015, Seattle, WA.
- Fresh Perspectives in Causal Inference II, *Joint Statistical Meetings*, 2014, Boston, MA.
- The Role of Causal Inference in Policy and Regulatory Decision Making, *Atlantic Causal Inference Conference*, 2013, Cambridge, MA.

- Fresh Perspectives in Causal Inference, *Joint Statistical Meetings*, 2013, Montreal, Canada.

Session organizer

- What is the Role of Causal Inference in Analyses of Electronic Medical Databases? *Atlantic Causal Inference Conference*, 2016, New York, NY.
- Fresh Perspectives in Causal Inference III, *Joint Statistical Meetings*, 2015, Seattle, WA.

Session chair, *Eastern North American Region/International Biometric Society Spring Meeting (ENAR)*, March, 2013. Orlando, FL.

Reviewer

- Grants

Mathematics and Statistics Discovery Grant, *Natural Sciences and Engineering Research Council of Canada*, 2017

Patient Centered Outcomes Research Institute, 2017, 2022.

Faculty grant committee 2016-2017, *Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute*

Postdoctoral Fellowship Award, *Fonds Wetenschappelijk Onderzoek - Vlaanderen, FWO*, 2015

- Conferences

32nd International Conference on Pharmacoepidemiology and Therapeutic Risk Management, 2016

The Society for Epidemiologic Research Annual Meeting, 2013

- Journals

Agency for Healthcare Research and Quality

Biometrics

Biostatistics

Biometrika

Clinical Trials

eGems

Epidemiology

Epidemiologic Methods

International Journal of Biostatistics

International Journal of Epidemiology

Pharmaceutical Statistics

PLOS Digital Health

American Journal of Epidemiology

Journal of Biopharmaceutical Statistics

Journal of Causal Inference

Journal of Clinical Epidemiology

Journal of Educational and Behavioral Statistics

Journal of the Royal Statistical Society

Journal of Statistical Software

Statistics and Probability Letters

Statistics in Medicine

Statistical Science Statistics in Biosciences

Teaching Statistics