Michele Santacatterina

Personal Data

EMAIL:	msanta@gwu.edu; msanta@bsc.gwu.edu
SKYPE ID:	miksanta86
PROFILES:	Personal Website, Google Scholar, LinkedIn, Research Gate
RESEARCH INTERESTS:	Biostatistics, Causal Inference, Data Science, Medicine

Academic Appointments

CURRENT 2020 AUG	Assistant Research Professor, Department of Biostatistics and Bioinformatics and Biostatistics Center, George Washington University School of Public Health

Education

Academic Degrees

2018 APR 2013 DEC	Ph.D., Biostatistics, Karolinska Institutet, Stockholm, Sweden Main topics: Biostatistics, Causal Inference, Data Science, Optimization.
2012 MAR 2009 OCT	M.S., Biostatistics, University of Milano-Bicocca, Milan, Italy Main topics: Classical and Bayesian inference, Survival analysis, Longitudinal data analysis, Statistical methods for Epidemiology, and Design of Experiments.
2009 SEP 2006 SEP	B.S., Statistics and Computer Science, University of Padua, Padua, Italy Selected topics: Probability, Statistical inference, Statistical models, Computational Statistics, Algebra, Linear Optimization, and Information Technology.

Relevant Courses

2016 MAR 2015 NOV	Course SF3961 "Statistical Inference" at KTH Royal Institute of Technology Organized and taught by Prof. Henrik HULT. 15-ECTS Ph.D.-level course within the Mathematical Statistics Ph.D. Program at KTH.
2015 JUN 2015 MAR	Course SF2822 "Applied nonlinear optimization" at KTH Royal Institute of Technology Organized and taught by Prof. Anders FORSGREN. 7.5-ECTS MSc-level course within the Applied Mathematics Program at KTH.

Other Education

2019 Jan   Auditing: [CS6780 - Advanced Machine Learning](#)
Taught by Prof. Thorsten Joachims, Cornell University, Department of
Computer Science and Department of Information Science.

2014 Dec   Course on ["Regression Models"](#) by John Hopkins University on [Coursera](#)
Organized and taught by Prof. Brian Caffo, Prof. Roger D. Peng and
Assoc. Prof. Jeff Leek, Johns Hopkins University.
[Certification and grades](#).

2013 Dec   Course on "Causal inference from observational data" at Karolinska
Institutet
Organized by the Swedish Interdisciplinary Graduate School at Karolinska Institutet. Taught by Prof. Miguel Hernán, Harvard School of Public
Health.

2013 Jun   Summer School in "Modern Methods in Biostatistics and Epidemiology", Treviso, Italy
Organized jointly by Harvard School of Public Health and Karolinska
Institutet.
Course attended:
"Causal Inference" by Prof. Andrea Rotnitzky, Harvard School of Public Health

2010 Feb   SAS Masterclass in Business Intelligence, Milan, Italy
Organized jointly by SAS and Accenture.
Certification obtained:
SAS Certified Base Programmer for SAS 9

2009 Aug   Visiting Student, Department of Mathematics, University of Barcelona
2009 Feb   Course attended: Probability, Statistical inference.

Research Experience

2018 Aug- 2020 Aug | Postdoctoral Associate, [TRIPODS Center for Data Science](#) and [Cornell Tech, Cornell University](#)
Development of novel statistical methods for optimal decision making using experimental and observational data with applications in public health and medicine.
Mentors: Prof. [Nathan Kallus](#), Prof. [Thorsten Joachims](#), Prof. [Kilian Weinberger](#)

2013 Dec - 2018 Apr | Ph.D. student, [Unit of Biostatistics, Karolinska Institutet](#)
Study and development of statistical methods based on mathematical programming techniques with application in public health and medical research. Teaching assistant.
Advisor: Prof. [Matteo Bottai](#)

2017 Jun - 2017 Sept | Visiting Scholar, [Cornell Tech, Cornell University](#)
Collaborate on research into novel optimization approaches to causal inference.
Faculty Sponsor: Prof. [Nathan Kallus](#)

2013 Jan - 2013 Nov	Research assistant, Unit of Biostatistics, Karolinska Institutet
	Application of statistical methods in public health research, with focus on HIV research. Writing of grant applications (KID funding). Teaching assistant.
2012 Mar - 2012 Dec	Research assistant, Department of Public Health, Karolinska Institutet
	Application of statistical methodologies for medical and public health research.
2011 Sept - 2012 Feb	Master's Thesis, Department of Public Health, Karolinska Institutet
	Title: Access to antiretroviral therapy and treatment effectiveness among injection drug users: results from a Swedish population based study.
	Advisor: Prof. Rino Bellocco
2008 Nov - 2009 Jan	Internship, Department of Pediatrics, University of Padua.
	Application and evaluation of methods for survival analysis using data from patients with anaplastic large-cell lymphoma.
	Contact persons: Prof. Laura Ventura and Dr. Gloria Tridello

Manuscripts and Publications

Work in progress

2020	*Kernel optimal orthogonality weighting: a balancing approach to estimating effects of continuous treatments* Kallus, N., and Santacatterina, M. (*)
	Novel optimization method for estimating causal dose-response curves.
2020	*Understanding and predicting the risk of bleeding disorders* Santacatterina, M. and Vaughan, R.
	Application of novel machine learning methods to help understand and predict the risk of bleeding disorders.
2020	*Optimal care for patients undergoing surgical interventions* Pennicooke, B., Santacatterina, M., Mummameni, P.
	Application of machine learning algorithms to identify clinical features and predict patient-reported outcomes among patients undergoing spine surgeries.
2020	*Quantile based regression for life expectancy estimation with censored data.* Garcia-Pareja, C., and Santacatterina, M., et. al.
	Novel statistical method for estimating life expectancy.

Statistical methodology

2019	*Optimal Weighting for Estimating Generalized Average Treatment Effects* Kallus, N., and Santacatterina, M. (*)
	Formulation of a new general causal estimand and development of an optimization method for its estimation.
2019	*CAB: Continuous Adaptive Blending estimator for policy evaluation and learning* Su, Y., Wang, L., Santacatterina, M., Joachims, T (NeurIPS 19, ICML 19)
	Novel method for estimating optimal targeted policies.

2019   *More robust estimation of sample average treatment effects using Kernel Optimal Matching in an observational study of spine surgical interventions* Kallus, N., Pennicooke, B., Santacatterina, M. (Under review in Statistics in Medicine) (*)

Extension and application of Kernel Optimal Matching to robustly estimate sample average treatment effects under lack of overlap.

2019   *Optimal balancing of time-dependent confounders for marginal structural models* Kallus, N., and Santacatterina, M. (*)

Novel method that estimates causal effects affected by time-dependent confounders from longitudinal observational studies.

2018   *Optimal probability weights for estimating causal effects of time-varying treatments with marginal structural Cox models* Santacatterina, M., Garcia-Pareja, C., Bellocco, R., Sönnerborg, A., Ekström, A.M., Bottai, M. (Statistics in Medicine)

Use of optimal probability weights when estimating treatment effects with longitudinal data. Comparisons, through simulations, with existing methods.

2018   *Optimal probability weights for inference with constrained precision* Santacatterina, M., and Bottai, M. (Journal of the American Statistical Association)

Proposal of a method for obtaining optimal probability weights based on solving a constrained nonlinear optimization problem.

(*) The order of authors follows alphabetical order.
Santacatterina, M. is the corresponding author of those manuscripts.

Healthcare policy and medical science

2020   *Does state malpractice environment affect outcomes following spinal fusions? A robust statistical and machine learning analysis of 549,775 discharges following spinal fusion surgery in the United States* Chan, A.K., Santacatterina, M., et.al. (Accepted to the Journal of Neurosurgery)

A medical study on the effect of malpractice claims on clinical outcomes.

2019   *The effect of age, fusion levels and their interactions on unfavorable outcomes and complications: A Retrospective Study of 60,000 patients* Pennicooke, B., Santacatterina, M., Elowitz, E., Kallus, N. (Submitted to Clinical Spine Surgery)

A medical study on the effect of surgical interventions, age and their interactions on clinical outcomes.

2017   *Risk behaviour determinants among people who inject drugs in Stockholm, Sweden over a 10-year period, from 2002 to 2012* Karlsson, N., Santacatterina, M., et. al. (Harm reduction journal)

A study on understanding the underlying causes that drive various types and levels of risk behaviours among people who inject drugs.

2017   *Effect of therapy switch on time to second-line antiretroviral treatment failure in HIV-infected patients* Häggblom, A., Santacatterina, M., et. al. (PloS one)

A medical study on the effect of therapy switch on time to second-line antiretroviral treatment failure among HIV-infected patients.

2016   *Impact of peer support on virologic failure in HIV-infected patients on antiretroviral therapy-a cluster randomized controlled trial in Vietnam* Sönnerborg, A., Tam, V.V., El-Khatib, Z., Santacatterina, M., et. al., (BMC infectious diseases)

A randomized control trial evaluating the impact of peer support on clinical outcomes in Vietnam.

2016 | *Inferences and conjectures in clinical trials: a systematic review of generalizability of study findings*   Santacatterina, M., and Bottai, M. (Journal of Internal Medicine)

Systematic review on generalizability and heterogeneity of study findings in randomized clinical trials.

2015 | *Correlates of mobile phone use in HIV care: Results from a cross-sectional study in South Africa*   Madhvani, N., Longinetti, E., Santacatterina, M., et. al. (Preventive medicine reports)

A cross-sectional study on the evaluation of the use of mobile phones as reminders in HIV care.

2014 | *Temporal Trends in the Swedish HIV-1 Epidemic: Increase in Non-B Subtypes and Recombinant Forms over Three Decades*   Neogi, U., Häggblom, A., Santacatterina, M., et. al. (PloS one)

A study on the temporal trend of the subtype distribution from the beginning of the HIV-1 epidemic in Sweden.

2014 | *The State-Led Large Scale Public Private Partnership 'Chiranjeevi Program' to Increase Access to Institutional Delivery among Poor Women in Gujarat, India: How Has It Done? What Can We Learn?*   De Costa, A., Kranti, S.V., Ryan, K., Sankara Raman, P., Santacatterina, M., et. al. (PloS one)

A study determining risk factors associated with disclosing HIV status among antiretroviral therapy recipients in South Africa.

2014 | *Gender perspective of risk factors associated with disclosure of HIV status, a cross-sectional study in Soweto, South Africa*   Longinetti, E., Santacatterina, M., El-Khatib, Z. (PloS one)

A cross-sectional study on risk factors for viremia and drug resistance.

2013 | *Antibiotic prescribing in women during and after delivery in a non-teaching, tertiary care hospital in Ujjain, India: a prospective cross-sectional study*   Sharma, M., Sanneving, L., Kalpana, M., Santacatterina, M., et.al. (Journal of Pharmaceutical Policy and Practice volume)

A study that aimed at present antibiotic prescribing among inpatients during and after delivery in rural India.

2012 | *Monitoring the efficacy of antiretroviral therapy by a simple reverse transcriptase assay in HIV-infected adults in rural Vietnam*   Cuong, D.D., Agneskog, E., Nguyen Thi Kim, C., Santacatterina, M., et.al. (Future medicine)

Determination of the feasibility of viral load monitoring in a cohort of HIV treatment-naive adult patients initiating antiretroviral therapy in rural Vietnam.

Conferences and Invited Talks

Organizer

Mar 14, 2021 | ENAR 2021

Organizer of the invited paper session: *Leveraging real-world data for improved medical decision-making: challenges, opportunities, and recent developments*

Dec 14, 2019 | NeurIPS, 2019
Organizer of the workshop: *"Do the right thing": machine learning and causal inference for improved decision making*, together with Thorsten Joachims (Cornell), Nathan Kallus (Cornell), Adith Swaminathan (Microsoft Research), David Sontag (MIT), and Angela Zhou (Cornell)

May 23, 2019 | Atlantic Causal Inference Conference 2019
Organizer of the symposium: *Optimization methods for causal inference*, together with Adam Kapelner (CUNY), Nathan Kallus (Cornell), Nikos Ignatiadis (Stanford), and Stefan Wager (Stanford)

Talks

Aug 03, 2020 | JSM 2020 - Health policy statistics section. *Optimal estimation of generalized average treatment effects using Kernel Optimal Matching.* Invited talk.

Mar 22, 2020 | ENAR 2020 - IMS invited session on challenges for precision medicine. *Kernel optimal orthogonality weighting: a balancing approach to estimating effects of continuous treatments.* Invited talk.

Jan 15, 2020 | McGill Department of Biostatistics - Biostatistics Seminar. *Kernel optimal orthogonality weighting: a balancing approach to estimating effects of continuous treatments.* Invited talk.

Sep 23, 2019 | Cornell Machine Learning in Medicine. *Kernel optimal orthogonality weighting: a balancing approach to estimating effects of continuous treatments.* Invited talk.

May 23, 2019 | Atlantic Causal Inference Conference 2019. *Optimal estimation of generalized average treatment effects using Kernel Optimal Matching.* Talk and poster.

Oct 24, 2018 | Second TRIPODS PI meeting. *Optimal balancing of time-dependent confounders for marginal structural models .* Talk.

Sep 7, 2018 | Cornell - AI Seminar. *Optimal Weighting for Causal Inference .* Invited talk.

Apr 11, 2018 | EUROCIM Causal Inference 2018. *Optimal balancing of time-dependent confounders for marginal structural models .* Talk.

Oct 18, 2017 | MELODEM Selection Group Meeting. *Optimal probability weights for inference with constrained precision .* Invited talk.

Sep 7, 2017 | Royal Statistical Society International Conference. *Estimating treatment effects with optimal inverse probability weighting.* Talk.

Aᴘʀɪʟ 5, 2017	**UK Causal Inference Meeting.** *Estimating treatment effects with optimal inverse probability weighting.* Poster.
Sᴇᴘ 7, 2016	**Royal Statistical Society International Conference.** *Optimal probability weights for inference with constrained precision .* Talk.
Sᴇᴘ 4, 2015	**Nordic and Baltic Stata Users Group meeting.** *Weight watchers: How to optimize your weight .* Invited talk.
Sᴇᴘ 28, 2014	**HIV Nordic conference.** *Antiretroviral therapy among HIV-infected people who inject drugs in Sweden: access and treatment response .* Poster.

Advising and Mentoring

Master Students

2019	**Yaniv Ravid**, M.Eng. in Operations Research and Information Engineering - Cornell Tech (Advisor). *A Python library for Kernel Optimal Matching*
2017	**Claudia Carlucci**, M.Sc. in Biostatistics - Univeristy of Milano-Bicocca and Karolinska Institutet (Mentor). *Modeling additive interaction with continuous variables: smoking and the risk for rheumatoid arthritis*
2015	**Chiara Chiavenna**, M.Sc. in Biostatistics - Univeristy of Milano-Bicocca and Karolinska Institutet (Mentor). *Laplace regression with censored data: an overview and application to observational data in cardiovascular epidemiology*

Tᴇᴀᴄʜɪɴɢ Exᴘᴇʀɪᴇɴᴄᴇ

2018 Aᴘʀ 2015 Aᴘʀ	Teaching Assistant at Karolinska Institutet
	Preparation of material and teaching of computer laboratory sessions.

- "Biostatistics I" for the Doctoral Program in Epidemiology,
- "Biostatistics II" for the Doctoral Program in Epidemiology,
- "Biostatistics I" for the Research School for Clinical Epidemiology, and,
- "Biostatistics II" for the Research School for Clinical Epidemiology.

The courses are given once per semester. They cover notions of statistics and biostatistics and the use of STATA for data analysis.

Fᴜɴᴅɪɴɢ

2017 Dᴇᴄ 2013 Dᴇᴄ	Karolinska Institutet Doctoral (KID) funding
	Project title: Novel methods for estimating optimal dynamic treatment regimes for HIV-infected patients. 1,280,000 SEK.

Written in collaboration with Prof. Matteo Bᴏᴛᴛᴀɪ (PI), Prof. Anna Mia Eᴋsᴛʀöᴍ, Prof. Anders Söɴɴᴇʀʙᴏʀɢ and Prof. Rino Bᴇʟʟᴏᴄᴄᴏ.

Other awards and merits

2017 Jun 2017 Aug	KI Travel Grant Visiting Scholar at Cornell Tech. 14,000 sek.
2012 Feb 2011 Sept	Programme Extra Plus - Fondazione Cariplo Visiting student with the aim of working on the Master's thesis at the Division of Global Health, Karolinska Institutet. 4,500 EUR.
2009 Aug 2009 Feb	Erasmus+ Grant Visiting student at the Department of Mathematics, University of Barcelona. 1,500 EUR.

SERVICE

Current 2019 Feb	Involved in the project WiTNY, Women in Technology and Entrepreneurship in New York.
Current 2017 Apr	Member of the Steering committee for the EURO-CIM
Current 2018 Aug	Reviewer for the *Journal of the Royal Statistical Society - Series A, Biometrical Journal, Computational Statistics and Data Analysis, NeurIPS, ICML, AISTAT, Nature Machine Intelligence, BMC Medical Research Methodology, Epidemiologicl methods, Journal of the American Statistical Association Statistics in Biopharmaceutical Research*

LANGUAGES

ENGLISH:	Fluent (C1/2).
ITALIAN:	Mother tongue.
SPANISH:	Basic knowledge (A2).
SWEDISH:	Basic knowledge (A2).

COMPUTER SKILLS

Operative Systems
Advanced user of Linux, MacOS and Windows.

Programming Languages

Advanced:	R and STATA
Intermediate:	Python, scikit-learn and SAS
Notions of:	Mathematica, SageMath, SQL, TensorFlow and Matlab

CHARACTER AND SKILLS

- Independent
- Cooperative
- Devoted
- Open-minded

OTHER INTERESTS

When I am not reading about statistics, I like to learn more about contemporary history. I also enjoy keeping myself active by playing sports and going to the gym. I especially enjoy running in the forest. It is my way of releasing stress and enjoy life. I love to play and listen to music, especially electronic

and film music. During the past few years I started several electronic music projects that included production using Ableton Live. Finally, together with my wife Alice, I have a dog named Mister Pom Princess "PK" Cake, which keeps me happily busy.