Session: Keynote (Wednesday 9th April, 9:15-10:15)

Location: Auditorium 1

Abstract author: Jonas Peters

Affiliation: ETH Zurich

Title: Hidden Confounders in Dynamical Systems

Abstract: Causal models can help us with predicting how a real-world system reacts under an active perturbation. Many causal methods and theoretical results have been developed for settings where data follow an i.i.d. structure. Often, however, data come from a dynamical system whose temporal structure cannot be ignored. In this talk, we argue that considering time-dependence does not only come with technical difficulties but also with benefits: we develop causal methods for dealing with hidden confounding that do not necessarily have a direct correspondence in the i.i.d. world. We show how they can be used for separating the effects of internal variability and external forcing in Earth system science, for example.

Session: Invited (Wednesday 9th April, 10:15-10:45)

Location: Auditorium 1

Abstract author: Erin Gabriel

Affiliation: University of Copenhagen

Title: Partial identification and ancillary covariates: What is the question?

Abstract: There has been a recent increase in the use of symbolic partial identification bounds for optimization of decision rules acknowledging that there may exist unmeasured confounders. Several works aim to manipulate existing sharp symbolic bounds to condition on and marginalize over ancillary covariates often with the desire to maintain sharp bounds. Making fewer assumptions about the ancillary covariates can make this easier, which highlights the conflicting nature of the assumptions one is willing to make and the ability to derive sharp partial identification bounds. In the extreme, with no assumptions or data, sharp bounds are the range of the parameter. To illuminate the differences in the assumptions made, I will give clear definitions of sharp and valid partial identification bounds with extensions of these definitions to include pointwise and uniform sharpness with the inclusion of ancillary covariates. I will give examples that demonstrate that when you have additional information about the covariate, sharpness can be hard to obtain, and I will discuss the potential for general criteria for obtaining sharp bounds in settings with ancillary covariates under what could be called full information.

Session: Sensitivity analysis and bounds (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 1

Abstract author: Alissa Gordon

Affiliation: University of California Berkeley

Title: Leveraging Covariates for Sensitivity Analysis of Hybrid Control Trials

Abstract: Many subpopulations cannot fully benefit from advances in health research due to lack of representation in trials. Barriers including limited funding and recruitment challenges cause this problem, leading to an evidence gap in research for these historically underrepresented groups. To bridge this gap, we propose a statistical methodology that allows researchers to safely use hybrid control trials. Hybrid control trials supplement traditional randomized controlled trials with external controls with the hopes of increasing trial efficiency. Current hybrid control trial estimators rely on a strict mean exchangeability assumption between the trial and external controls, introducing bias when violated. Unfortunately, this assumption is impractical in the sense that it is almost always violated.To mitigate against this bias, we leverage a non-parametric sensitivity analysis that reliably bounds the bias. Unlike others, this sensitivity analysis does not impose any additional assumptions on the data, provides easy interpretation, and uses debiased machine learning for efficient estimation. This sensitivity analysis uses omitted variable bias methodologies and covariates to explain the violations in this mean exchangeability assumption. With its flexibility and simplicity, researchers will be able to make informed conclusions after comparing the strength of evidence against the size of the potential bias. Simulations have confirmed that the sensitivity analysis reliably bounds the bias and that the additional control units increase efficiency even after accounting for bias. By safely using historical controls in addition to concurrent trial controls, our methodology will allow for reduced sample sizes, budgets, and timelines, making trials focusing on marginalized, underfunded, or small populations more feasible. Ultimately, this approach aims to increase the volume of research dedicated to historically underrepresented groups, contributing to greater future health equity.

Additional Authors: Alejandro Schuler

Session: Sensitivity analysis and bounds (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 1

Abstract author: Sam Pimentel

Affiliation: University of California, Berkeley

Title: Choosing an estimand for weighted observational studies using design sensitivity

Abstract: Choices about observational study design, notably the choice of estimand, have important implications for whether the final estimate will exhibit robustness to unmeasured confounding. In practice however, the aspects of a study that influence sensitivity to unmeasured confounding are not well understood or accounted for when planning a study. We demonstrate how design sensitivity, a quantity describing the asymptotic power of a sensitivity analysis, can be used to compare multiple candidate estimands in weighted observational studies to improve robustness to unmeasured bias. Specifically, using data from a referendum on the 2016 Colombian peace agreement we explore how altering the definition of treatment and altering the target population of interest impact the expected performance of sensitivity analyses. We also consider how the choice of a specific method of sensitivity analysis interacts with the choice of estimand to influence robustness to unmeasured confounding.

Additional Authors: Melody Huang, Dan Soriano

Session: Sensitivity analysis and bounds (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 1

Abstract author: Jannis Kueck

Affiliation: Heinrich Heine University, Düsseldorf Institute for Competition Economics (DICE)

Title: Sensitivity Analysis for Difference-in-Differences Models

Abstract: Difference-in-Differences models are widely used for policy evaluation in empirical research, relying on the parallel trend assumption for identification of causal parameters. Parallel trends are assumed to hold either unconditionally or conditionally on observable pre-treatment covariates. However, the potential presence of unobserved pre-treatment or time-varying covariates challenges this identification. This paper introduces a framework for sensitivity analysis of key causal quantities in Difference-in-Differences models. We evaluate the robustness of treatment effect estimates in a multi-period design against violations of identification assumptions and propose relevant diagnostics for empirical applications. We demonstrate the usefulness of our method in a simulation study and empirical examples. Finally, our study provides evidence on machine-learning based estimation.

Additional Authors: Philipp Bach, Sven Klaassen, Mara Mattes, Martin Spindler

Session: Sensitivity analysis and bounds (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 1

Abstract author: Luke Keele

Affiliation: UPenn

Title: Local Effects of Continuous Instruments without Positivity

Abstract: Instrumental variables are a popular study design for the estimation of treatment effects in the presence of unobserved confounders. In the canonical instrumental variables design, the instrument is a binary variable. In many settings, however, the instrument is continuous. Standard estimation methods can be applied with continuous instruments, but they require strong assumptions. While recent work has introduced more flexible estimation approaches, these methods require a positivity assumption that is implausible in many applications. We derive a novel family of causal estimands using stochastic dynamic interventions that allows a range of intervention distributions that are continuous with respect to the observed distribution of the instrument. These estimands focus on a specific local effect but do not require a positivity assumption. Next, we develop doubly robust estimators for these estimands that allow for estimation of the nuisance functions via nonparametric estimators. We use empirical process theory and sample splitting to derive asymptotic properties of the proposed estimators under weak conditions. In addition, we derive methods for profiling the principal strata as well as a method of sensitivity analysis. We evaluate our methods via simulation and demonstrate their feasibility using an application on the effectiveness of surgery for specific emergency conditions.

Additional Authors: Alex Levis, Prabrisha Rakshit

Session: Interference and continuous exposures (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 2

Abstract author: Fabrizia Mealli

Affiliation: European University Institute

Title: Bipartite Interference: Causal Inference when Intervention Units and Outcome Units Differ

Abstract: We study causal inference in settings characterized by interference with a bipartite structure. There are two distinct sets of units: intervention units to which an intervention could be applied and outcome units on which the outcome of interest can be measured. Outcome units may be affected by interventions on some, but not all, intervention units, as captured by a bipartite graph.Examples of this setting can be found across many applications, for example in analyses of the impact of pollution abatement in plants on health outcomes for individuals, or the effect of transportation network expansions on regional economic activity.Using these examples, we introduce and discuss a variety of causal estimands for these bipartite settings that are associated with actionable policies and stochastic interventions.We discuss how knowledge of the bipartite graph allows unbiased estimation of these causal quantities in rather general experimental settings and the threat of nontrivial positivity violations.

Additional Authors: Laura Forastiere, Guido Imbens, Georgia Papadogeorgou, Zhaoyan Song, Cory Zigler

Session: Interference and continuous exposures (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 2

Abstract author: Vanessa McNealis

Affiliation: University of Glasgow

Title: Revisiting the effects of maternal education on adolescents' academic performance: Doubly robust estimation in a network-based observational study

Abstract: In many contexts, particularly when study subjects are adolescents, peer effects can invalidate typical statistical requirements in the data. For instance, it is plausible that a student's academic performance is influenced both by their own mother's educational level as well as that of their peers. Since the underlying social network is measured, the Add Health study provides a unique opportunity to examine the impact of maternal college education on adolescent school performance, both direct and indirect. However, causal inference on populations embedded in social networks poses technical challenges, since the typical no interference assumption no longer holds. While inverse probability-of-treatment weighted (IPW) estimators have been developed for this setting, they are often highly unstable. Motivated by the question of maternal education, we propose doubly robust (DR) estimators combining models for treatment and outcome that are consistent and asymptotically normal if either model is correctly specified. In this work, we assume that the network is a union of connected subnetworks and propose doubly robust (DR) estimators combining models for treatment and outcome that are consistent and asymptotically normal if either model is correctly specified. We present empirical results that illustrate the DR property and the efficiency gain of DR over IPW estimators even when the treatment model is misspecified. Simulations are conducted under different scenarios of (latent) treatment dependence. Contrary to previous studies, our robust analysis does not provide evidence of an indirect effect of maternal education on academic performance within adolescents' social circles in Add Health.

Additional Authors: Erica E. M. Moodie, Nema Dean

Session: Interference and continuous exposures (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 2

Abstract author: Kaitlyn Lee

Affiliation: UC Berkeley

Title: Bridging Binarization: Causal Inference with Dichotomized Continuous Exposures

Abstract: The average treatment effect (ATE) is a common parameter estimated in causal inference literature, but it is only defined for binary exposures. Thus, despite concerns raised by some researchers, many studies seeking to estimate the causal effect of a continuous exposure create a new binary exposure variable by dichotomizing the continuous values into two categories. In this paper, we affirm binarization as a statistically valid method for answering causal questions about continuous exposures by showing the equivalence between the binarized ATE and the difference in the average outcomes of two specific modified treatment policies. These policies impose cut-offs corresponding to the binarized exposure variable and assume preservation of relative self-selection. Relative self-selection is the ratio of the probability density of an individual having an exposure equal to one value of the continuous exposure variable versus another. The policies assume that, for any two values of the exposure variable with non-zero probability density after the cut-off, this ratio will remain unchanged. Through this equivalence, we clarify the assumptions underlying binarization and discuss how to properly interpret the resulting estimator. Additionally, we introduce a new target parameter that can be computed after binarization that considers the status-quo world. We argue that this parameter addresses more relevant causal questions than the traditional binarized ATE parameter. Finally, we present a simulation study to illustrate the implications of these assumptions when analyzing data and to demonstrate how to correctly implement estimators of the parameters discussed.

Additional Authors: Alan Hubbard, Alejandro Schuler

Session: Interference and continuous exposures (Wednesday 9th April, 11:15-12:35)

Location: Auditorium 2

Abstract author: Anja Shahu

Affiliation: Columbia University Mailman School of Public Health

Title: Estimating effects of longitudinal modified treatment policies (LMTPs) on rates of change in health outcomes with repeated measures data

Abstract: Longitudinal modified treatment policies (LMTPs) quantify the effects of interventions that depend on the natural value of exposure, generalizing "stochastic" and "shift" interventions as well as other policy-relevant quantities. The current LMTP estimation approach yields effects on outcomes measured at the end of a study; however, repeated measures data often contains time-varying outcomes measured at each visit and interest may lie in estimating effects on the rate of change in these outcomes over time. For example, one may wish to quantify the effect of an LMTP on the rate of progression of a disease. We extend the LMTP approach to estimate the effect on change in a time-varying outcome over time and propose a hypothesis testing framework to formally test whether there is a difference in change in the outcome over time under an LMTP versus the natural outcome trajectory (or versus a different LMTP). Repeated measures data also frequently has unique data complications that must be considered. One such complication is that of irregular visit times, where the visit timing varies among individuals from some pre-specified time. We propose an extension to our work that permits effect estimation and hypothesis testing for an LMTP in a setting with irregular visit times. We present results from a simulation study which shows that ignoring irregular visit times may lead to bias, and we illustrate our hypothesis testing framework in both regular and irregular visit time settings.

Additional Authors: Daniel Malinsky

Session: Randomized studies (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 1

Abstract author: Lin Liu

Affiliation: Shanghai Jiao Tong University

Title: Covariate adjustment in RCT motivated by Higher-Order Influence Functions

Abstract: Higher-Order Influence Functions (HOIF), developed in a series of papers over the past twenty years, is a fundamental theoretical device for constructing rate-optimal causal-effect estimators from observational studies. However, the value of HOIF for analyzing well-conducted randomized controlled trials (RCTs) has not been explicitly explored. In the recent U.S. Food and Drug Administration (FDA) and European Medicines Agency (EMA) guidelines on the practice of covariate adjustment in analyzing RCTs, in addition to the simple, unadjusted difference-in-mean estimator, it was also recommended to report the estimator adjusting for baseline covariates via a simple parametric working model, such as a linear model. In this paper, we show that a HOIF-motivated estimator for the treatment-specific mean has significantly improved statistical properties compared to popular adjusted estimators in practice when the number of baseline covariates p is relatively large compared to the sample size n. We also characterize the conditions under which the HOIF-motivated estimator improves upon the unadjusted one. Furthermore, we demonstrate that a novel debiased adjusted estimator proposed recently by Lu et al. is, in fact, another HOIF-motivated estimator in disguise. Numerical and empirical studies are conducted to corroborate our theoretical findings.

Additional Authors: Sihui Zhao, Xinbo Wang, Xin Zhang

Session: Randomized studies (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 1

Abstract author: Wenxin Zhang

Affiliation: University of California Berkeley

Title: Towards Efficient Statistical Inference and Optimal Design in Adaptive Experiments

Abstract: Adaptive experiments play a crucial role in clinical trials and online A/B testing. Unlike static, non-adaptive trial designs, adaptive experimental designs dynamically adjust treatment randomization probabilities and other key design elements in response to data collected sequentially during the experiment. These designs are useful for achieving different objectives, such as reducing uncertainty in causal estimand estimation or improving benefit of participants within the experiment. Despite their advantages, the adaptive nature of these designs and the time-dependent nature of the data introduce significant challenges in making unbiased statistical inferences from non-i.i.d. data.Building upon the Targeted Maximum Likelihood Estimator (TMLE) literature that has provided valid statistical inference tailored to adaptive experimental settings using inverse weighting strategies tailored for adaptive experiment settings, we propose a new TMLE that improves the efficiency for estimating causal estimands under adaptive designs. Beyond efficient statistical inference, we further introduce a general framework for implementing optimal adaptive designs, customized to achieve various objectives efficiently. The performance of our proposed estimators and adaptive designs is demonstrated through theoretical analysis and extensive simulations.

Additional Authors: Mark van der Laan

Session: Randomized studies (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 1

Abstract author: Mathieu Even

Affiliation: Inria

Title: Rethinking the Win Ratio: A Causal Framework for Hierarchical Outcome Analysis

Abstract: Quantifying causal effects in the presence of complex and multivariate outcomes is a key challenge to evaluate treatment effects. For hierarchical multivarariates outcomes, the FDA recommends the Win-Ratio and Generalized Pairwise Comparisons approaches (Pocock, 2011; Buyse, 2010). However, as far as we know, these empirical methods lack causal or statistical foundations to justify their broader use in recent studies. To address this gap, we establish causal foundations for hierarchical comparison methods.We define related causal effect measures, and highlight that depending on the methodology used to compute Win Ratios or Net Benefits of treatments, the causal estimand targeted can be different, as proved by our consistency results. Quite dramatically, it appears that the causal estimand related to the historical estimation approach can yield reversed and incorrect treatment recommendations in heterogeneous populations, as we illustrate through striking examples. In order to compensate for this fallacy, we introduce a novel, individual-level yet identifiable causal effect measure that better approximates the ideal, non-identifiable individual-level estimand. We prove that computing Win Ratio or Net Benefits using a nearest-neighbor pairing approach between treated and controlled patients, an approach that can be seen as an extreme form of stratification, leads to estimating this new causal estimand measure. We extend our methods to observational settings via propensity weighting, distributional regression to address the curse of dimensionality, and a doubly robust framework. We prove the consistency of our methods, and the double robustness of our augmented estimator. These methods are straightforward to implement, making them accessible to practitioners. Finally, we validate our approach using synthetic data, the AIDS dataset, and the CANTO observational oncology dataset.

Additional Authors: Julie Josse

Session: Randomized studies (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 1

Abstract author: Georgi Baklicharov

Affiliation: Ghent University

Title: Nearly Assumption-Free Methods for Treatment Effect Testing in Randomized Trials with Intercurrent Events

Abstract: Intercurrent events, such as treatment switching, rescue treatment, and truncation by death, pose significant challenges to the interpretation of treatment effects in randomized clinical trials. Intention-to-treat analyses often fail under these conditions, potentially resulting in misleading conclusions. Existing methods, including hypothetical estimands and survivor average causal effects, address some challenges but rely on strong assumptions, are prone to positivity violations, and struggle with time-varying confounders.In this talk, we present a novel methodology for analyzing longitudinal clinical trial data impacted by intercurrent events. Our approach does not require data on time-varying confounders and does not exclude positivity violations on the intercurrent events. It relies on a weak structural assumption about the occurrence of intercurrent events and is found to deliver only small bias under its violation. We propose asymptotically efficient, model-free tests of the null hypothesis of no treatment effect, which make use of data-adaptive nuisance parameter estimates. In the context of randomized experiments, we moreover propose asymptotically efficient tests in a subclass of tests that have greater robustness properties. The methodology's empirical performance is demonstrated through simulation studies and the re-analysis of a recent diabetes trial, which is complicated by truncation due to death.

Additional Authors: Kelly Van Lancker, Stijn Vansteelandt

Session: Randomized studies (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 1

Abstract author: Ellen Hamaker

Affiliation: Methodology and Statistics, Faculty of Social Sciences, Utrecht University

Title: The average causal effect: Across people, across time, and across disciplines

Abstract: In psychology and related disciplines, a persistent question is how findings derived from comparisons between individuals at a single time point relate to insights gained from comparisons across multiple time points within a single individual. Although clear statistical and substantive explanations for obtaining different results exist, the issue has not been thoroughly addressed from a causal inference perspective yet.In this presentation, I will begin by outlining the discussion as it appears in the psychological literature, and then formalize it within a causal framework. The point of departure for this is the individual causal effect, which is the difference between two potential outcomes for a specific individual at a specific time point. From there, we can average either across individuals at the same time point yielding the person average causal effect (PACE), or across time points within the same person yielding the time average causal effect (TACE).Using panel data from a micro-randomization study—where multiple individuals are measured at multiple time points with randomized exposure levels for each person and time point—I will cover the usual identification assumptions necessary to link the PACE to data from a randomized controlled trial. I will then investigate how—and under what conditions—similar assumptions can be formulated for the TACE, and connect this with recent work on potential outcomes and time series experiments. By integrating ideas and insights from various disciplines and presenting these within a single framework, I aim to identify how they relate and complement each other.

Session: Time-to-event data (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 2

Abstract author: Murthy Mittinty

Affiliation: College of Medicine and Public Health, Flinders Health and Medical Research Institute, Flinders University , South Australia, Australia

Title: Doubly robust estimation of relative survival in the presence of informative censoring

Abstract: In cancer research relative survival is a common method to estimate the survival of patients. Relative survival, which is the ratio of the observed survival to expected survival, is used when the cause of death is unknown. In the estimation of Relative survival when the age of diagnosis changes the administrative censoring due to study end may be informative. This is a well-known problem in the literature and methods exist for accounting for informative censoring. Usually, in the absence of informative censoring marginal relative survival rates are used to study the cancer population prognosis. The difference in marginal survival rates, along with some assumptions, between exposed and unexposed groups is used for studying the causal effects. Syriopoulou et al (2021) have developed the IPTW and doubly robust standardised methods in relative survival framework to obtain the average causal effects. The methods developed in this manuscript does not consider informative censoring. In this paper we show an extension of this method that accounts for informative censoring. A real-world data on multiple primary cancers is provided and a simulation study is conducted to investigate how sensitive are the methods to model misspecification and censoring misspecification, including different ways for obtaining the variance and standard errors.

Session: Time-to-event data (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 2

Abstract author: Nan van Geloven

Affiliation: Leiden University Medical Center

Title: Doubly robust estimation of marginal cumulative incidence curves for competing risk analysis

Abstract: Competing risk scenarios arise when individuals can experience multiple event types and the occurrence of one event type precludes occurrence of the other(s). A common goal is to estimate the cumulative incidence, that is, the probability of a specific event happening over time accounting for the possibility that other events may occur beforehand. This talk focuses on estimating marginal cumulative incidence curves under different treatment conditions when using observational data challenged by baseline confounding. Potential approaches to address this include using an inverse probability of treatment weighted (IPW) Aalen-Johansen estimator or using outcome modeling, for instance by standardizing over cause-specific hazard regressions. For these two approaches to work, correct specification of either the treatment model or of the cause-specific hazards models is needed respectively. We propose a novel augmented IPW estimator that combines the treatment and outcome models. A key advantage of this approach is its "doubly robust" property, meaning it provides consistent estimates if at least one of the two models is correctly specified. This extends a previously introduced doubly robust estimator for standard survival data (Wang 2018) to the setting with competing events. The method utilizes pseudo-observations to handle censored data, under the assumption of a completely independent censoring process. Through simulations, the performance of these three adjustment methods is evaluated in finite sample settings, including scenarios with model misspecification for the treatment-covariates and outcomes-covariates relationships. The practical application of these methods is demonstrated in a cohort study of breast cancer patients, estimating marginal cumulative incidence curves for recurrence, second primary tumor development and death following mastectomy or breast-conserving therapy. Wang J. A simple, doubly robust, efficient estimator for survival functions using pseudo observations. Pharm Stat. 2018;17(1):38-48. doi:10.1002/pst.1834

Additional Authors: Patrick van Hage, Saskia le Cessie, Marissa C. van Maaren, Hein Putter

Session: Time-to-event data (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 2

Abstract author: Mark Knudsen

Affiliation: Section of Biostatistics, University of Copenhagen

Title: The Markov property and causal inference for time-varying treatments using the Cox model

Abstract: When using the Cox model to analyze the effect of a time-varying treatment on a survival outcome, treatment is commonly included using only the current level as a time-dependent covariate. Such a model does not necessarily assume that past treatment has no effect on the outcome (the Markov property), since it is possible to model the hazard conditional on only the current treatment value. But modeling the hazard conditional on the full treatment history is required in order to interpret the results causally, and such a full model assumes the Markov property when only including current treatment. Relying on the Markov property is problematic since one would generally expect it not to hold in settings with unmeasured frailty. We show why this is the case, even if the true causal effect of treatment really only depends on its current value. Further, we provide an example of a correctly specified Cox model, including only current treatment as a covariate, where the Markov property is not fulfilled. Transforming the result to survival scale does not give the true intervention-specific survival probabilities, showing that analysts should be careful to include sufficient information about treatment history when making causal statements from such models.

Additional Authors: Erin Gabriel, Torben Martinussen, Helene Rytgaard, Arvid Sjölander

Session: Time-to-event data (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 2

Abstract author: Torben Martinussen

Affiliation: Section of Biostatistics, University of Copenhagen

Title: Assumption-lean variable significance testing in a survival analysis context.

Abstract: The difficult and subtle interpretation of hazard ratios has sparked growing criticism, prompting researchers to explore alternative treatment effect estimands for time-to-event endpoints. This evolution has itself evoked concerns, in view of inherent weaknesses affecting those alternatives and, therefore, hazard ratios still attract some interest.However, corresponding tests and hazard ratio estimates fall short due to their reliance on models, potentially leading to biased test results and an inadequate mapping of survival risks in case of model misspecification. To address some of these limitations, we present in this talk some testing procedures that are based on debiased machine learning strategies.

Additional Authors: Stijn Vansteelandt

Session: Time-to-event data (Wednesday 9th April, 14:15-15.55)

Location: Auditorium 2

Abstract author: Matthew Pryce

Affiliation: London School of Hygiene and Tropical Medicine

Title: Estimating heterogeneous causal contrasts using infinite dimensional targeting in time to event data

Abstract: In recent years, there has been a growing interest in causal machine learning (ML) estimators which can estimate heterogenous treatment effects using time to event (TTE) data. TTE data is frequently collected in medical, engineering, and financial settings, with its temporal dimension allowing researchers to explore treatment effect heterogeneity over time. However, due to this temporal dimension, estimating treatment effect heterogeneity can be challenging, with observations being either right censored (leaving the study prior to an event being observed), or left truncated (entering the study at different times). A range of causal ML estimators have been proposed which focus on overcoming these issues, with notable examples being those that attenuate regularisation bias induced via estimated nuisance parameters, including causal survival forests, which handle right censored data, and more recently the ltrc-DR/ltrc-R estimators, which handle left truncation as well as right censoring. These estimators have been shown to hold favourable oracle efficiency properties, however, each of these estimator either fails to account for left truncation or requires the estimation of challenging nuisance functions. Additionally, none of these estimators fully harness the temporal structure of the treatment effect, requiring the effect be estimated at each time point separately, leaving them prone to producing non-smooth treatment effects over time. In this work we aim to address these issues, presenting surv-iTMLE, an alternative influence function-based approach for estimating heterogenous causal contrasts using TTE data. surv-iTLME uses a two-step approach to estimate heterogenous causal contrasts, including the difference and ratio in conditional survival probability between binary exposure groups. It uses infinite dimensional targeted learning (iTMLE) to generate de-biased treatment effect predictions at time point of interest, regressing these against the covariates of interest to produce a smooth treatment effect curve over time. We present surv-iTMLE's favourable convergence and robustness properties through simulated data examples, highlighting its performance in comparison with existing causal ML estimators for TTE data. We also demonstrate its application by reviewing treatment effect heterogeneity between immunotherapy initiators and patients who receive chemotherapy amongst patients with non-small cell lung cancer.

Additional Authors: Karla Diaz-Ordaz, Ruth Keogh, Stijn Vansteelandt

Session: Data fusion (Wednesday 9th April, 16:25-17.25)

Location: Auditorium 1

Abstract author: Vanessa Rodriguez

Affiliation: UCL

Title: Causal machine learning for generalizing heterogeneous treatment effects

Abstract: Randomized Controlled Trials (RCTs), although considered as the gold standard for causal effect estimation, may suffer from limited external validity. This can stem from factors such as sample unrepresentativeness due to narrow eligibility criteria or differential response to trial participation within the population. Consequently, treatment decisions may favour individuals closely resembling the average participant, at the expense of underrepresented groups. This issue is exacerbated in the presence of treatment effect heterogeneity. A variety of approaches are available to generalize inferences from a randomized trial to a target population, with much of the existing literature focused on the average treatment effect (ATE). However, the relative performance of methods for generalizing conditional average treatment effects (CATEs) using machine learning (ML) meta-learners remains under explored, particularly under varying degrees of sampling bias, CATE complexity, and runtime confounding. Moreover, obtaining valid inference using such ML models poses significant challenges. In this talk, we evaluate two methods: a T-Learner, which uses two separate models to learn the conditional mean potential outcomes and inverse probability of sampling weights (IPSW) but lacks rate robustness when using ML models, and a generalized DR-Learner, a debiased machine learning estimator that addresses these limitations. We present an in-depth comparison of these methods for generalizing CATEs from a trial nested in the target population of interest. Following an extensive simulation study, we observe that the generalised DR-Learner consistently exhibits lower median mean squared error (MSE) than both the standard and generalized T-Learner in almost all cases, but especially in settings with complex sampling mechanisms and smaller sample sizes. We also explore the use of conformal prediction to ensure valid inference, which has traditionally been a limitation of using ML methods for causal inference. We present the coverage and interval lengths obtained by using weighted conformal inference, allowing us to obtain prediction intervals for causal effects under covariate shift. We anticipate these findings will provide practical guidance to practitioners wanting to incorporate ML methods in their analysis. A variety of approaches are available to generalise inferences from a randomized trial to a target population, with much of the existing literature focused on the average treatment effect (ATE). However, the relative performance of methods for generalising conditional average treatment effects (CATEs) using machine learning (ML) meta-learners remains under explored, particularly under varying degrees of sampling bias, CATE complexity, and runtime confounding. Moreover, obtaining valid inference using such ML models poses significant challenges. In this talk, we evaluate two methods: a T-Learner,

which uses two separate models to learn the conditional mean potential outcomes and inverse probability of sampling weights (IPSW) but lacks rate robustness when using ML models, and a generalized DR-Learner, a debiased machine learning estimator that addresses these limitations. We present an in-depth comparison of these methods for generalising CATEs from a trial nested in the target population of interest. Following an extensive simulation study, we observe that the generalised DR-Learner consistently exhibits lower median mean squared error (MSE) than both the standard and generalized T-Learner in almost all cases, but especially in settings with complex sampling mechanisms and smaller sample sizes. We also explore the use of conformal prediction to ensure valid inference, which has traditionally been a limitation of using ML methods for causal inference. We present the coverage and interval lengths obtained by using weighted conformal inference, allowing us to obtain prediction intervals for causal effects under covariate shift. We anticipate these findings will provide practical guidance to practitioners wanting to incorporate ML methods in their analysis.

Additional Authors: Karla Diaz-Ordaz, Brieuc Lehmann

Session: Data fusion (Wednesday 9th April, 16:25-17.25)

Location: Auditorium 1

Abstract author: Pan Zhao

Affiliation: University of Cambridge

Title: Causal mediation analysis of data fusion with application to bridging risk and relative efficacy of vaccines

Abstract: Refined vaccine regimens with variant-matched inserts are routinely approved by the regulatory agencies based on historical phase 3 clinical trials and immunobridging studies. Historical phase 3 clinical trials often help establish immune biomarkers that can reliably predict the risk or vaccine efficacy (VE) against a clinical endpoint. Once one or more immune correlates have been established, an immunobridging study, rather than another VE trial, will be conducted to compare the immunogenicity of an updated vaccine against that of an approved vaccine. In this article, we develop efficient and robust statistical methods that estimate the relative vaccine efficacy (relVE) of an updated vaccine versus an approved vaccine against the currently circulating strain, using relevant patient-level historical trials and immunobridging data. We discuss in detail identification assumptions, propose efficient and multiply robust estimators, and evaluate the finite sample performance of our proposed estimators. We demonstrate meaningful efficiency gain, which would translate to a smaller sample size when designing an immunobridging study, using our proposed estimators. We applied our framework to estimating the relative VE of multiple bivalent mRNA-1273 vaccines against monovalent prototype mRNA-1273 vaccine using data from the COVID-19 Variant Immunologic Landscape (COVAIL) Trial.

Additional Authors: Oliver Dukes, Bo Zhang

Session: Data fusion (Wednesday 9th April, 16:25-17.25)

Location: Auditorium 1

Abstract author: Rickard Karlsson

Affiliation: Delft University of Technology

Title: Robust integration of external control data in randomized trials

Abstract: In this talk, we consider methods for improving the efficiency of randomized trials through the use of "external controls" – individuals who received the control treatment in the trial during routine practice or in prior experimental studies. Existing external control methods, however, can have substantial bias if the populations underlying the trial and the external control data are not exchangeable. Here, we characterize a randomization-aware class of treatment effect estimators in the population underlying the trial that remain consistent and asymptotically normal when using external control data, even when exchangeability does not hold. We consider two members of this class of estimators: the well-known augmented inverse probability weighting trial-only estimator, which is the efficient estimator when only trial data are used; and a more efficient member of the class when exchangeability holds and external control data are available, which we refer to as the optimized randomization-aware estimator. To achieve robust integration of external control data in trial analyses, we then propose a combined estimator based on the efficient trial-only estimator and the optimized randomization-aware estimator. We show that the combined estimator is consistent and no less efficient than the most efficient of the two component estimators, whether the exchangeability assumption holds or not. We examine the estimators' performance in simulations and we illustrate their use with data from two trials of paliperidone extended-release for schizophrenia.

Additional Authors: Guanbo Wang, Piersilvio De Bartolomeis, Jesse H. Krijthe, Issa J. Dahabreh

Session: Matching and weighting (Wednesday 9th April, 16:25-17.25)

Location: Auditorium 2

Abstract author: Jaehyuk Jang

Affiliation: Department of Statistics, Seoul National University

Title: Mixing Samples to Address Weak Overlap in Causal Inference

Abstract: In observational studies, the assumption of sufficient overlap (positivity) is fundamental for the identification and estimation of causal effects. Failing to account for this assumption yields inaccurate and potentially infeasible estimators. To address this issue, we introduce a simple yet novel approach, Mixing, which mitigates overlap violations by constructing a synthetic treated group that combines treated and control units. Our strategy offers three key advantages. First, it improves estimator accuracy by preserving unbiasedness while reducing variance. The benefit is particularly significant in settings with weak overlap, though the method remains effective regardless of the overlap level. This phenomenon results from the shrinkage of propensity scores in the mixed sample, which enhances robustness to poor overlap. Second, it enables direct estimation of the target estimand without discarding extreme observations or modifying the target population, thus facilitating straightforward interpretation of the results. Third, the mixing approach is highly adaptable to various weighting schemes, including contemporary methods such as Entropy Balancing. The estimation of the Mixed IPW (MIPW) estimator is done via M-estimation, and the method extends to a broader class of weighting estimators through a resampling algorithm. We illustrate the mixing approach through extensive simulation studies and provide practical guidance with a real-data analysis.

Additional Authors: Suehyun Kim, Kwonsang Lee

Session: Matching and weighting (Wednesday 9th April, 16:25-17.25)

Location: Auditorium 2

Abstract author: Sharon-Lise Normand

Affiliation: Harvard Medical School

Title: Causal Inference Approaches to Assessing Healthcare Providers

Abstract: While causal inference approaches are increasingly common, their application in assessing the quality of healthcare providers is limited and often challenging. Complications arise due to small provider sample sizes, low outcome event rates, and between-center heterogeneity in patient case-mix. The de facto strategy employs parametric regression modeling using indirect standardization, outside the causal framework, invokes strong assumptions of linearity in the covariates, and makes limited use of a large pool of confounders. Focusing on the average treatment effect on the treated, the typical policy estimand, we use constrained optimization algorithms to balance confounders and reweight outcomes to obtain counterfactual estimates of surgical center quality. Our approach provides a flexible framework with robust estimation of nuisance functions that accommodates between-center heterogeneity in both the types of surgeries provided and patient confounders. Covariate balancing propensity score weights, entropy balancing weights, and stable balancing weights are compared against standard methods in the policy literature.

This work is funded by Grant R01HL162893 from the U.S. National Institutes of Health.

Additional Authors: Larry Han

Session: Matching and weighting (Wednesday 9th April, 16:25-17.25)

Location: Auditorium 2

Abstract author: Jose Zubizarreta

Affiliation: Harvard University

Title: Weighting for Personalized and Sample-Bounded Meta Analyses

Abstract: The field of causal inference has experienced rapid growth over the past few decades, with substantial advancements in identification, estimation, and inference. While much of this progress has been confined to individual studies, our understanding of causes and effects ultimately hinges on the integration, reconciliation, and synthesis of multiple study designs and data sources. We present a weighting methodology for the synthesis of evidence derived from a series of separate studies, also known as meta-analysis. This methodology encompasses the traditional fixed- and random-effect approaches for meta-analyses and generalizes them. We use this approach to personalize meta-analyses, targeting the covariate profiles of individuals of interest in a sample-bounded manner, thereby enhancing their personalization and robustness. We propose a technique to detect studies that significantly deviated from the defined profile, suggesting when it might be prudent to exclude them from the analysis. We establish multiple consistency conditions and demonstrate asymptotic normality for the proposed estimator. We illustrate this approach through an empirical study.

Additional Authors: Wenqi Shi

Session: Keynote (Thursday 10th April, 9:00-10:00)

Location: Auditorium 1

Abstract author: Vasilis Syrgkanis

Affiliation: Stanford University

Title: Detecting clinician implicit biases in diagnoses using proximal causal inference and debiased machine learning

Abstract: Clinical decisions to treat and diagnose patients are affected by implicit biases formed by racism, ableism, sexism, and other stereotypes. These biases reflect broader systemic discrimination in healthcare and risk marginalizing already disadvantaged groups. Existing methods for measuring implicit biases require controlled randomized testing and only capture individual attitudes rather than outcomes. However, the "big-data" revolution has led to the availability of large observational medical datasets, like EHRs and biobanks, that provide the opportunity to investigate discrepancies in patient health outcomes. In this work, we propose a causal inference approach to detect the effect of clinician implicit biases on patient outcomes in large-scale medical data. Specifically, our method uses proximal mediation to disentangle pathway-specific effects of a patient's sociodemographic attribute on a clinician's diagnosis decision. We test our method on real-world data from the UK Biobank. Our work can serve as a tool that initiates conversation and brings awareness to unequal health outcomes caused by implicit biases.

Additional Authors: Kara Liu, Russ Altman

Session: Invited (Thursday 10th April, 10:00-10.30)

Location: Auditorium 1

Abstract author: Stéphanie van der Pas

Affiliation: Vrije Universiteit Amsterdam

Title: Regression discontinuity and feedback between design and analysis

Abstract: The key quantity in the regression discontinuity design (RDD) is the cutoff on the 'running variable'. Treatment gets assigned based on someone's value of the running variable in relation to the cutoff. Here we propose new methods for data where a cutoff was used, but its value is unknown. We find that feedback between the 'design' and 'analysis' stage is beneficial. We discuss this in the context of other, primarily propensity-score based methods, where efforts are made to avoid feedback.

Additional Authors: Julia Kowalska, Mark van de Wiel

Session: Causal machine learning (Thursday 10th April, 11:00-11.40)

Location: Auditorium 1

Abstract author: Nima Hejazi

Affiliation: Harvard Chan School of Public Health

Title: Evaluating the effects of continuous exposures under interference with induced modified treatment policy

Abstract: Many contemporary scientific domains, such as environmental health and epidemiology, feature observational data whose structure poses several distinct challenges for causal inference. Distinct problems arising in such contexts include the presence of continuous exposures; an often large number of putative confounders of the causal relationships of interest; and network interference, wherein the exposure of one unit may affect the outcome of its neighbors. Each of these problems requires care to address. We introduce a novel intervention scheme, the induced modified treatment policy, designed to aid in identification of the causal effect attributable to intervening on a continuous exposure while simultaneously addressing network interference between study units; this addresses the first and third of the challenges noted above. Building on recent theoretical developments, we demonstrate how to estimate the causal effects of induced modified treatment policies, developing two types of asymptotically (semi-parametric) efficient estimators that are both compatible with the use of nonparametric regression and/or machine learning techniques for initial estimation of nuisance functions; this addresses the second challenge noted above by allowing for the analyst to flexibly adjust for possibly many confounders. In numerical experiments, we illustrate how the induced modified treatment policy eliminates causal (i.e., identification) bias that arises from interference in the network setting. We apply the new methodology to evaluate the causal effect of zero-emission vehicle uptake on air pollution in California, strengthening evidence from prior analytic studies.

Additional Authors: Salvador Balkus

Session: Causal machine learning (Thursday 10th April, 11:00-11.40)

Location: Auditorium 1

Abstract author: Martin Spindler

Affiliation: University of Hamburg

Title: DoubleMLDeep: Estimation of Causal Effects with Multimodal Data

Abstract: This paper explores the use of unstructured, multimodal data, namely text and images, in causal inference and treatment effect estimation. We propose a neural network architecture that is adapted to the double machine learning (DML) framework, specifically the partially linear model. An additional contribution of our paper is a new method to generate a semi-synthetic dataset which can be used to evaluate the performance of causal effect estimation in the presence of text and images as confounders. The proposed methods and architectures are evaluated on the semi-synthetic dataset and compared to standard approaches, highlighting the potential benefit of using text and images directly in causal studies. Our findings have implications for researchers and practitioners in economics, marketing, finance, medicine and data science in general who are interested in estimating causal quantities using non-traditional data.

Additional Authors: Sven Klaassen, Jan Teichert-Kluge, Philipp Bach, Victor Chernozhukov, Suhas Vijaykumar

Session: Fairness (Thursday 10th April, 11:00-11.40)

Location: Auditorium 2

Abstract author: Razieh Nabi

Affiliation: Emory University

Title: Statistical learning for constrained functional parameters in infinite-dimensional models with applications in fair machine learning

Abstract: Constrained learning has become increasingly important in machine learning, especially in algorithmic fairness where predictive models are specifically designed to meet pre-defined fairness criteria. This work studies constrained statistical learning from a statistical functional perspective. It focuses on estimating a function-valued parameter of interest, characterized as the minimizer of a risk criterion, under constraints where one or more fairness-related real-valued parameters are set to zero or bounded. This talk particularly focuses on counterfactual and causal constraints, which have emerged as an important framework for quantifying fairness notions. Often, closed-form solutions exist for the optimal functional parameter under causal constraints, offering insight into the mechanisms that enforce fairness in predictive models. Results also suggest natural estimators for the constrained parameter, which can be derived by combining estimates of unconstrained parameters from the data-generating distribution. As a result, fair machine-learning algorithms can be implemented seamlessly alongside any statistical learning method or off-the-shelf software.

Additional Authors: David Benkeser, Nima Hejazi, Mark van der Laan

Session: Fairness (Thursday 10th April, 11:00-11.40)

Location: Auditorium 2

Abstract author: Youmi Suk

Affiliation: Columbia University

Title: Rethinking Item Fairness with Counterfactuals and Single-World Intervention Graphs

Abstract: Since the 1960s, the testing community has strived to ensure fair assessment practices. Differential item functioning (DIF) is a widely used statistical notion for assessing items that may unfairly disadvantage specific subgroups of test-takers (e.g., females). However, traditional DIF analyses focus only on statistical relationships in observed data and cannot explain why such unfairness occurs. To address this limitation, we introduce a novel counterfactual framework for defining and detecting unfair items using single-world intervention graphs (SWIGs). By leveraging SWIGs and potential (i.e., counterfactual) outcomes, we define counterfactual DIF as the difference in item functioning between two hypothetical worlds: one where an individual belongs to one subgroup and another where they belong to a different subgroup, while holding their ability constant. We also connect counterfactual DIF to related fairness concepts, including group versus individual fairness, item impact, and intersectionality. In particular, we use SWIGs to graphically distinguish between item fairness at the population level and the individual level. Additionally, we discuss causal identification strategies within the SWIG framework and provide a statistical method based on outcome regression for detecting causally unfair items. The performance of our approach is evaluated through simulation studies across various fairness scenarios by comparing it to traditional DIF detection methods. We further demonstrate its practical application using real data from the Programme for International Student Assessment (PISA). Finally, we discuss the broader implications of promoting causal fairness in testing and assessment practices.

Additional Authors: Weicong Lyu

Session: Heterogeneous treatment effects (Thursday 10th April, 11:45-12.25)

Location: Auditorium 1

Abstract author: Zijun Gao

Affiliation: University of Southern California

Title: Selective inference for data-driven subgroups based on biomarkers

Abstract: In randomized experiments with heterogeneous treatment effects, subgroup analysis provides significant benefits, such as personalized treatment recommendations, but poses challenges for inference when subgroups are learned from data. Motivated by the German Breast Cancer Study, where subgroups are defined using a biomarker threshold—a common practice in clinical trials—we develop a design-based inference procedure tailored to this type of subgroup selection.The validity of our method relies solely on knowledge of the randomization mechanism, requiring no assumptions about the underlying model, making it particularly suitable for complex datasets. Compared to sample-splitting based inference, our approach is deterministic and avoids the power loss associated with reduced sample size for inference. The computation of our method is often similar to a standard randomization test without selection and requires no intricate sampling procedures to approximate conditional distributions. Furthermore, when predefined biomarkers are unavailable, we extend the method by incorporating a data-driven biomarker while maintaining the desirable properties of the approach. We demonstrate the validity and efficiency of our methods through the analysis of the GBCS dataset and simulated data.

Session: Heterogeneous treatment effects (Thursday 10th April, 11:45-12.25)

Location: Auditorium 1

Abstract author: Richard Post

Affiliation: Erasmus Medical Center

Title: Beyond Conditional Averages: Estimating The Individual Causal Effect Distribution

Abstract: In recent years, the field of causal inference from observational data has emerged rapidly. The literature has focused on (conditional) average causal effect estimation. However, when (remaining) variability of individual causal effects (ICEs) is considerable, (conditional) average effects may be uninformative for an individual. Deriving the ICE distribution requires knowledge of the joint distribution of potential outcomes, but the latter is unidentifiable without making assumptions due to the fundamental problem of causal inference. In this talk, I show that the ICE distribution is identifiable under conditional independence of the individual effect and the potential outcome under no exposure, in addition to the common assumptions of consistency, positivity, and conditional exchangeability. Moreover, I present a family of flexible latent variable models that may be used to study individual effect modification and estimate the ICE distribution from cross-sectional data. How such latent variable models may be applied and validated in practice is illustrated in a case study on the effect of Hepatic Steatosis on a clinical precursor to heart failure. Under the assumptions presented, I estimate that the effect for 20.6% (95% Bayesian credible interval: 8.9%, 33.6%) of the population is greater than twice the harmful average causal effect. Finally, I will discuss the difficulty of assessing the identifiability assumption in practice based on subject matter knowledge.

Additional Authors: Edwin van den Heuvel

Session: Meta-analysis (Thursday 10th April, 11:45-12.25)

Location: Auditorium 2

Abstract author: Qingyang Shi

Affiliation: University of Groningen

Title: Causally-interpretable meta-analysis using aggregate data

Abstract: Evidence synthesis and meta-analysis inform clinical practice guidelines and health economic evaluations. However, heterogeneity of treatment effects poses a significant challenge. Standard meta-analysis addresses this through a random-effects distribution, but fails to explicitly link treatment effects to patient characteristics, limiting interpretation in specific target populations. Meta-regression explores heterogeneity by regressing treatment effects on study-level mean covariates, but is subject to ecological bias and high variance due to limited between-trial variability. A causally-interpretable meta-analysis offers a more rigorous framework with causal estimands and assumptions and provides several estimators using individual patient data (IPD). However, IPD is often unavailable in practice. This paper proposes a new approach towards causally-interpretable meta-analysis using only aggregate data, addressing the limitations of both traditional methods and IPD-based approaches. Our method leverages reported marginal and subgroup-specific treatment effects, along with summary statistics of patient characteristics, to estimate heterogeneous treatment effects. The average treatment effect in the target population is obtained by the plug-in estimator. This approach also offers an alternative to standard meta-regression by incorporating subgroup treatment effects to address ecological bias and reduce variance. We illustrate the application of our method using a meta-analysis of SGLT2 inhibitors in heart failure patients. A new R package `CIMAgD` is developed for implementation.

Additional Authors: Wouter van Amsterdam; Sacha la Bastide van Gemert; Veerle Coupé; Talitha Feenstra; Issa Dahabreh

Session: Meta-analysis (Thursday 10th April, 11:45-12.25)

Location: Auditorium 2

Abstract author: Tat Thang Vo

Affiliation: University Paris XII

Title: Causally interpretable meta-analysis under restricted access to individual-level data

Abstract: Obtaining causally interpretable meta-analysis results is challenging when there are differences in the distribution of effect modifiers between eligible trials. To overcome this, recent work on transportability methods has considered standardizing results of individual studies over the case-mix of a target population, prior to pooling them as in a classical random-effect meta-analysis. One practical challenge, however, is that case-mix standardization often requires the individual participant data (IPD) to be fully accessible in every eligible studies and in the target population, which is difficult due to privacy concerns. In this paper, we aim to develop novel strategies to integrate aggregated-level data from eligible trials with non-accessible IPD into a causal meta-analysis, via extending moment-based methods frequently used for population-adjusted indirect comparison in health technology assessment. Since valid inference for these moment-based methods by M-estimation theory requires additional aggregated data that are often unavailable in practice, computational methods to address this concern are also developed. We assess the finite-sample performance of the proposed approaches by simulated data, and then apply these approaches on real-world clinical data to investigate the effectiveness of risankizumab versus ustekinumab among patients with moderate to severe psoriasis.

Additional Authors: Stijn Vansteelandt, Shu Yang, Antoine Chambaz

Session: Time-varying confounding and dynamic treatment regimens (Thursday 10th April, 14:05-15.45)

Location: Auditorium 1

Abstract author: Xi Lin

Affiliation: University of Oxford

Title: Simulating Longitudinal Data from Marginal Structural Models

Abstract: Simulating longitudinal data from specified Marginal Structural Models (MSMs) is important for evaluating causal inference methods and designing clinical trials. However, this task is challenging. While data generation typically proceeds in a fully conditional manner using structural equations according to a temporal ordering, MSMs require capturing causal effects that are marginal over time-dependent confounders, making it difficult to align conditional distributions with target marginal quantities. We propose a flexible and efficient algorithm for simulating longitudinal data that adheres exactly to a specified MSM. Given the importance of time-to-event outcomes in clinical trials, we extend the method to handle survival MSMs. Compared to existing approaches, our method offers several key advantages. It enables exact simulation from a known MSM rather than relying on approximations, avoids imposing restrictive assumptions on the data-generating process, and is efficient as it only requires evaluating analytic functions instead of computationally intensive techniques such as Monte Carlo approximations or numerical integration. We validate the algorithm's efficacy through extensive simulation studies replicating real-world scenarios. Our method will facilitate researchers in effectively simulating data with target causal structures for their specific scenarios.

Additional Authors: Robin J. Evans

Session: Time-varying confounding and dynamic treatment regimens (Thursday 10th April, 14:05-15.45)

Location: Auditorium 1

Abstract author: Alec McClean

Affiliation: Division of Biostatistics NYU Grossman School of Medicine

Title: Longitudinal trimmed treatment effects

Abstract: Propensity score trimming addresses positivity violations by excluding individuals with extreme scores, but existing methods typically focus on single-timepoint data. We extend trimming to longitudinal settings and connect it to stochastic dynamic interventions — another approach for addressing positivity violations that modifies interventional propensity scores. We first demonstrate that trimmed effects based on cumulative cross-world propensity scores correspond to differences in mean potential outcomes under simultaneous interventions across all timepoints. These effects may lack interpretability, so we propose a complementary alternative based on longitudinal modified treatment policies (LMTPs); these interventions depend only on covariate history and possibly the natural value of treatment. Differences in mean potential outcomes under these LMTPs yield interpretable longitudinal trimmed effects. We illustrate their efficacy with simple examples. To enable efficient estimation, we develop smooth approximations of all considered trimmed effects and show that they remain definable as differences in mean potential outcomes under stochastic dynamic interventions. After establishing identification, we derive efficient influence functions for each parameter and construct sequentially doubly robust estimators that achieve root-n consistency and asymptotic normality under nonparametric conditions. We demonstrate the practical application of our methods through an analysis of longitudinal ICU data.

Additional Authors: Ivan Diaz

Session: Time-varying confounding and dynamic treatment regimens (Thursday 10th April, 14:05-15.45)

Location: Auditorium 1

Abstract author: Ignacio Gonzalez Perez

Affiliation: EPFL

Title: A new encoding of time-varying treatments and its implications for causal inference

Abstract: We describe an alternative encoding of time-varying treatment strategies. This encoding is designed to facilitate the representation of identifiability assumptions that can be represented in causal Directed Acyclic Graphs (DAGs) and Single World Intervention Graphs (SWIGs), while giving the identified quantities a per-protocol interpretation. We show how this encoding can simplify the exposition of conventional identifiability assumptions, and derive new results for interventionist analysis in settings where the treatment of interest is time-varying. Furthermore, we propose several estimators, including one that is semi-parametrically efficient and doubly robust. These results are illustrated through an analysis of the Systolic Blood Pressure Intervention Trial (SPRINT), where we, for the first time, give a per-protocol interpretation to its findings.

Additional Authors: Mats J. Stensrud

Session: Time-varying confounding and dynamic treatment regimens (Thursday 10th April, 14:05-15.45)

Location: Auditorium 1

Abstract author: Chunyu Wang

Affiliation: Zhejiang University

Title: Model selection in estimating optimal dynamic treatment regimes

Abstract: Optimal dynamic treatment regimes (DTRs), as a key part of precision medicine,have gained more and more attention recently. To inform clinical decision making,interpretable and parsimonious models for contrast functions are preferred, whichinevitably to some extent suffer from misspecification. It is therefore important toproperly evaluate the performance of the candidate interpretable models and select onewith better approximation to the unknown contrast function. Moreover, as a DTRusually involves multiple decision points, an inaccurate approximation at a latterdecision point will then affect its estimation at an earlier point when a backwardinduction algorithm is applied. This paper aims to perform model selection for contrastfunctions in the context of learning optimal DTRs from observed data. Note that therelative performance of candidate models may heavily depend on the sample size when,for example, the comparision is made between parametric models and tree-based models.Therefore, instead of investigating the limit behavior of each candidate model anddeveloping methods to asymptotically select the correct one, we focus on the finitesample performance of each model and attempt to perform model selection under a givensample size. To this end, we adopt the counterfactual cross-validation metric andpropose a novel method to estimate the variance of the metric. Supplementing thecross-validation metric with its estimated variance allows us to characterize theuncertainty on model selection under a give sample size and facilitates conducting ahypothesis test over a preferred model structure. Simulation studies are provided todemonstrate (i) the performance of our proposed variance estimator and (ii) theimprovement achieved by incorporating model selection for contrast functions inestimating optimal DTRs. We finally apply our method to the analysis of SequentialTreatment Alternatives to Relieve Depression (STAR*D) data.

Additional Authors: Brian Tom

Session: Time-varying confounding and dynamic treatment regimens (Thursday 10th April, 14:05-15.45)

Location: Auditorium 1

Abstract author: Catharina Stoltenberg

Affiliation: University of Oslo

Title: Causal effects of adding opioid-sparing medication to opioid treatments on subsequent opioid use

Abstract: Patients who use opioids for acute pain management face an increased risk of developing substance dependence. Finding opioid-sparing treatment alternatives is therefore important to reduce opioid misuse. We consider an intervention, referred to as the add-on regime, where opioid-sparing medications, such as nonsteroidal anti-inflammatory drugs (NSAIDs), are always prescribed concomitantly with opioids. This is a dynamic treatment which depends on the natural value of treatment. The regime assigns NSAIDs at a given time if and only if a physician would have prescribed either opioids or NSAIDs at that time. The add-on regime is more aligned with standard prescription practices than conventional static and dynamic regimes. We define average causal effects of add-on regimes on subsequent opioid use and derive sufficient conditions for their identification. We show that these conditions, under certain assumptions, are weaker than established conditions, and prove two identification formulas consistent with existing results. We emphasize that the presented theory is general, in the sense that it applies to a any setting where one treatment is given concurrently with another. The theoretical results are applied to unique data from a cohort of Norwegian trauma patients, linking the Norwegian National Trauma Registry to several national databases, including the Norwegian Prescription Database recording all drugs dispensed by prescription in Norway.

Additional Authors: Matias Janvin, Jon Michael Gran

Session: Graphical models and causal discovery (Thursday 10th April, 14:05-15.45)

Location: Auditorium 2

Abstract author: Nadja Rutsch

Affiliation: Vrije Universiteit Amsterdam

Title: Causal Inference in Finite Samples: The Potential of Invalid Adjustments

Abstract: Traditional covariate selection methods for causal inference focus on achieving unbiasedness and asymptotic efficiency. In many practical scenarios, researchers must estimate causal effects from observational data with limited sample sizes or when covariates are difficult or costly to measure. Their needs might be better met by selecting adjustment sets that are finite sample-optimal in terms of Mean Squared Error (MSE). We aim to find the adjustment set that minimizes the MSE of the causal effect estimator, taking into account the joint distribution of the variables and the sample size. We present examples where the MSE-optimal adjustment set differs from the asymptotically optimal adjustment set. To identify the MSE-optimal adjustment set, we introduce a sample size criterion for comparing adjustment sets in linear Gaussian models. We develop graphical criteria to reduce the search space for this adjustment set based on the causal graph. In experiments with simulated data, we show that the MSE-optimal adjustment set can outperform the asymptotically optimal adjustment set in finite sample size settings, making causal inference more practical in such scenarios.

Additional Authors: Stéphanie van der Pas, Sara Magliacane

Session: Graphical models and causal discovery (Thursday 10th April, 14:05-15.45)

Location: Auditorium 2

Abstract author: Elizabeth Ogburn

Affiliation: Johns Hopkins University

Title: Missing data with causal and statistical dependence

Abstract: Despite the growing interest in causal and statistical inference for settings with data dependence, few methods currently exist to account for missing data in dependent data settings; most classical missing data methods in statistics and causal inference treat data units as independent and identically distributed (i.i.d.). We introduce the notion of entangled missingness, which occurs when missingness indicators may exhibit causal or statistical dependence across units. We propose graphical models to represent entangled missingness and derive sound and complete identification results for the full data distribution. Next we consider the special case of a bivariate outcome under univariate treatment assignment and with missingness in both components of the outcome. This may represent paired subjects (e.g. parent and child pairs) or bivariate outcomes associated with a single subject (e.g. a risk and a benefit associated with treatment). We derive influence function-based estimators under missingness at random and block conditional missingness.

Additional Authors: Rohit Bhattacharya, Razieh Nabi, Dan Scharfstein, Ilya Shpitser, Ranjani Srinivasan, Shiyao Xu

Session: Graphical models and causal discovery (Thursday 10th April, 14:05-15.45)

Location: Auditorium 2

Abstract author: Kai Teh

Affiliation: UCL

Title: Causal Interpretation of Anterial Graphs and Constrained Confounder Selection

Abstract: We provide a valid causal interpretation for anterial graphs, a class of graphs containing directed acyclic graphs, that are also closed under conditioning and marginalisation. In this setting, we provide a graphical procedure that returns a graph which is valid in jointly representing post-intervened variables and pre-intervened (observational) variables, thus extending single world intervention graphs, originally introduced by Richardson and Robins (2013). Using this graphical representation, we provide an element-wise procedure of selecting confounders given flexibly prescribed set constraints.

Additional Authors: Kayvan Sadeghi, Terry Soo

Session: Graphical models and causal discovery (Thursday 10th April, 14:05-15.45)

Location: Auditorium 2

Abstract author: Hyunseung Kang

Affiliation: University of Wisconsin-Madison

Title: A Causal Inference Perspective on Single-Cell Perturb-Seq Experiments

Abstract: The integration of functional genomics, CRISPR-Cas9, and single-cell technologies, like scRNA-seq, has advanced studies on the effects of genetic perturbations on molecular phenotypes at the single cell level, particularly through single-cell Perturb-seq experiments. Briefly, single-cell Perturb-seq experiments use CRISPR-based technologies to silence (or activate) a specific, target gene in a cell (i.e., the study unit) and multiple gene expressions between the "treated" cells (i.e., cells where the target gene was silenced) and the "control" cells (i.e., cells where the target gene was not silenced) are compared. Despite rapid progress in the design of Perturb-seq experiments over the past decade, analysis of data from these experiments is mostly associational and lacks a formal causal framework, notably (a) a precise, causal definition of the scientific quantities of interest from these experiments, (b) plausibility (or implausibility) of identifying assumptions from the experiments, and critically, (c) a formal framework for defining potential sources of measured and unmeasured confounding. This work addresses this gap by highlighting two causal frameworks for Perturb-seq experiments: a causal mediation analysis with unmeasured mediator-outcome confounding and discovery of gene regulatory network (GRNs) with a novel Poisson-log normal (PLN) model. The proposed causal mediation analysis framework identifies mediating genes and their direct and mediated effects on transcription factors (TF), even in the presence of unmeasured confounders affecting TFs. The proposed causal discovery algorithm is useful to identify directed regulatory pathways from count-based, overdispersed transcriptomic data, a hallmark feature of single-cell Perturb-seq data. Notably, unlike most existing methods for discovering GRN based on observational expressional data, we show that the estimates from our PLN-based framework are less sensitive to biases from unmeasured confounders due to the design of Perturb-seq experiments. To assess the usefulness of the proposed causal frameworks in practice, we validate our results through complementary epigenomic and proteomic evidence. ChIP-seq datasets confirm the presence of regulatory interactions at key promoter and enhancer regions, and protein–protein interaction networks that carry functional relationships among regulators and their targets. These multi-modal, biological validations underscore the robustness of causal frameworks to reveal intricate regulatory circuitry and guide future therapeutic discoveries.

Additional Authors: Jingqi Duan, Zhongxuan Sun, Sunduz Keles

Session: Graphical models and causal discovery (Thursday 10th April, 14:05-15.45)

Location: Auditorium 2

Abstract author: Anne Helby Petersen

Affiliation: University of Copenhagen

Title: Interpretable evaluation of causal discovery algorithms: Are we doing better than random guessing?

Abstract: Causal discovery algorithms intend to recover (parts) of causal data generating mechanisms by analyzing empirical data they generated. A highly active research community has produced a plethora of different algorithms. Many have well-established properties in the large sample limit, but differ strongly when applied to real (finite) data. Their finite sample properties are typically evaluated using simulations and a few select real data examples with known data generating mechanisms. However, there does not exist a general guideline for how such evaluation studies should be designed, and therefore, comparing results across different studies is difficult. As a result, we cannot say what algorithms should be most suited for what tasks, which is a large obstacle for applying causal discovery in practice. We propose to establish a simple common evaluation baseline for causal discovery algorithms by posing the question: How much better than random guessing are we doing? Perhaps surprisingly, the answer is sometimes "not significantly better", even for well-established algorithms. We address the question by providing mathematical insights into the expected behavior of using random guessing for a specific subtask of causal discovery, namely skeleton estimation (i.e. estimating a DAG without orienting the edges). We derive exact distributional results for a range of typical causal discovery evaluation metrics (including precision and recall) under random guessing and show that in certain scenarios, these metrics are prone to achieve very large values simply by chance. We also propose an exact test of overall skeleton fit, and showcase its use on a real data application. Finally, we suggest a general simulation-based pipeline for using random guessing negative controls for more general causal discovery evaluation metrics, and showcase its use.Overall, we find that casual discovery is not just another machine learning problem: Estimating a high-dimensional object such as a DAG is difficult, and evaluating how well one did is equally challenging. We believe that reporting results from random guessing negative controls will be a useful next step towards more transparent and interpretable evaluations, and ultimately, a better understanding of what we can hope to learn by applying causal discovery in practice.

Session: Invited (Thursday 10th April, 16:15-16:45)

Location: Auditorium 1

Abstract author: Linbo Wang

Affiliation: University of Toronto

Title: The revival of relative survival

Abstract: The hazard ratio is the most commonly reported measure in survival analysis, yet it suffers from well-documented limitations, particularly when used for causal interpretation. These include its time-varying nature and the selection bias inherent in period-specific hazard ratios. In this talk, I revisit relative survival and cumulative incidence ratios as alternative measures that directly compare marginal survival/failure probabilities and offer more robust causal interpretations. I introduce a novel modeling framework that targets relative survival functions and constructs coherent nuisance models to ensure identifiability and interpretability. This approach leads to a class of flexible, covariate-adjusted estimators that avoid strong parametric assumptions, accommodate effect heterogeneity, and enable the estimation of adjusted survival curves. Applications to clinical trial data illustrate the practical advantages of this framework over hazard-based methods, especially in settings where rare events and dynamic treatment effects complicate traditional analyses.

Session: Invited (Thursday 10th April, 16:45-17:15)

Location: Auditorium 1

Abstract author: Julia Rohrer

Affiliation: Leipzig University

Title: How Can We Make Rigorous Causal Inference More Mainstream?

Abstract: Correlation does not imply causation—but a narrow focus on hammering this catchphrase home to applied researchers may have left many of them ill-equipped to tackle the causal inference problems they will inevitably encounter in their work. In psychology (but also in other fields), researchers often fail to realize that they are asking a causal question in the first place, and even if they do, they may lack the toolkit to arrive at a coherent answer. How can we as a scientific community do better? In my talk, I will discuss various (hopefully) helpful ways forward, with a special focus on what the more technically inclined can do to help applied researchers—potentially afraid of anything related to statistics—improve their inferences.

Session: Transportability and target trials (Friday 11th April, 09:00-10:40)

Location: Auditorium 1

Abstract author: Daniel de Vassimon Manela

Affiliation: University of Oxford

Title: Testing Generalizability with Frugal Parameterization

Abstract: Ensuring robust model performance across diverse real-world scenarios requires addressing both transportability across domains with covariate shifts and extrapolation beyond observed data ranges. However, there is no formal procedure for statistically evaluating generalizability in machine learning algorithms, particularly in causal inference. Existing methods often rely on arbitrary metrics like AUC or MSE and focus predominantly on toy datasets, providing limited insights into real-world applicability. To address this gap, we propose a systematic and quantitative framework for evaluating model generalizability under covariate distribution shifts, specifically within causal inference settings. Our approach leverages the frugal parameterization, allowing for flexible simulations from fully and semi-synthetic benchmarks, offering comprehensive evaluations for both mean and distributional regression methods. By basing simulations on real data, our method ensures more realistic evaluations, which is often missing in current work relying on simplified datasets. Furthermore, using simulations and statistical testing, our framework is robust and avoids over-reliance on conventional metrics. Grounded in real-world data, it provides realistic insights into model performance, bridging the gap between synthetic evaluations and practical applications. We also introduce a non-parametric generative model that fits the frugal parameterization model on real data, and equivalence testing as our test method.

Additional Authors: Linying Yang, Xinwei Shen, Robin Evans

Session: Transportability and target trials (Friday 11th April, 09:00-10:40)

Location: Auditorium 1

Abstract author: Mats Stensrud

Affiliation: EPFL

Title: An Overlooked Stability Property of the Risk Ratio and Its Practical Implications

Abstract: Risk ratios are widely used effect measures in empirical research, but their stability and transportability across populations remain debated. Here, we show that the causal risk ratio is stable under selection based on immune status. For example, the causal risk ratio remains unchanged when individuals who cannot experience the outcome, regardless of treatment, are excluded from a study. We term this property "immune-selection stability" (ISS).ISS applies broadly and generalizes previous findings on the stability of risk ratios. In particular, it includes the properties proposed by Huitfeldt and colleagues as special cases. Furthermore, unlike earlier results, ISS does not rely on assumptions about cross-world counterfactuals. We also demonstrate an analogous property for survival ratios.Despite decades of discussion on the properties of risk ratios, ISS has received little to no attention. However, its implications for interpreting, comparing, and transporting estimates across populations are considerable. We illustrate the practical relevance of ISS by assessing the results of an HIV vaccine trial and a fertility treatment study.

Additional Authors: Marco Piccininni

Session: Transportability and target trials (Friday 11th April, 09:00-10:40)

Location: Auditorium 1

Abstract author: Rhian Daniel

Affiliation: Cardiff University

Title: Transportability from first principles via regression by composition

Abstract: Most work on generalizability / transportability has focused on identifying a set of measured variables conditional on which the outcome of interest and 'population' can be assumed to be independent. ('Population' here could refer to a binary random variable taking value 0 for the population from which the study units are sampled and 1 for the target population of interest to which the 'transportation' of results is desired.)Finding such a conditional independence assumption to be unrealistic in many settings, Anders Huitfeldt and colleagues (Epidemiology 34(3), 396-399; arXiv:2106.06316) instead propose a first-principles mechanistic approach, which can in some settings lead to uncovering particular effect measures that can be shown to be transportable under the proposed mechanistic model. Building on a recent and related paper ('Does God toss logistic coins?' and other questions motivating Regression by Composition, JRSS-A 187(3), 636-655), in this presentation, we take the famous ACTG175 trial (of combination antiretroviral therapies vs. mono-therapies for HIV patients) as a case study and consider the transportation of the results from the relatively healthy predominantly white male U.S. study population to an Ethiopian cohort with a higher distribution of baseline risks. We propose a first-principles causal model to explain the heterogeneity in the treatment effect when comparing didanosine and zidovudine in combination against zidovudine alone on death or disease progression within 2 years of randomization. Such a mechanistic model includes latent variables such as whether the outcome would occur under no medication, whether a normal and sufficient metabolic response to each drug would occur if taken, and whether the side effects of the two drugs in combination would be tolerated. By considering plausible relationships between CD4 count (the most influential baseline covariate) and each of the latent variables, we exhibit the features of plausible L'Abbé plots (risk transformations) for the treatment effect curve across the full range of baseline risks. Whilst these do not correspond to risk transformations arising from standard (e.g. GLM) models, we show how they emerge from the more flexible 'regression by composition' framework, and compare our results to those from more standard approaches to transportability.

Additional Authors: Daniel Farewell

Session: Transportability and target trials (Friday 11th April, 09:00-10:40)

Location: Auditorium 1

Abstract author: Lorenzo Gasparollo

Affiliation: EPFL

Title: An overlooked bias in target trial emulations and how to fix it

Abstract: Many datasets involve staggered entries, where individuals join the study at different points in time. For example, randomized controlled trials (RCTs) in medicine usually recruit patients over time, and electronic health records contain information from the time a patient enters the healthcare system. Seminal works by Hernán, Brumback, and Robins (2000) and Hernan et al. (2008) on target trials used such datasets, treating the time an individual entered the study as a covariate in regression models.In this talk, I will describe a subtle – but frequently overlooked – positivity violation that appears in the analysis of staggered entry data. Because of this positivity violation, a frequently used class of Inverse Probability Weighting Censoring procedures leads to biased results.I will then propose new adjustment methods to circumvent such bias, and elaborate on how these fit with the target trial emulation framework.Finally, I will outline and compare these two approaches in settings wherein the bias varies in severity, thereby clarifying when one method is preferred over the other.

Additional Authors: Mats Stensrud

Session: Transportability and target trials (Friday 11th April, 09:00-10:40)

Location: Auditorium 1

Abstract author: Sebastien Haneuse

Affiliation: Harvard T.H. Chan School of Public Health

Title: Robust causal inference for point exposures in EHR-based target trial emulations with missing information on study eligibility

Abstract: The target trial emulation framework has emerged an important tool for the conduct of observational comparative effectiveness studies based on large scale electronic health record (EHR) databases. Central to the framework is that it requires analysts to formalize key aspects of the "hypothetical trial", including eligibility criteria. Missingness in variables that define eligibility criteria, however, is a pervasive challenge in EHR-based settings and it is typically the case that patients with incomplete eligibility information are excluded from analysis without consideration of assumptions that are being made (implicitly), leaving the study conclusions subject to potential selection bias. To the best of our knowledge, however, very little work has been done to mitigate this concern, and existing solutions require correct specification of all relevant models to ensure consistent estimation of causal contrasts. In this work, we propose a robust and efficient estimator of the causal average treatment effect on the treated study eligible population in cohort studies where eligibility defining covariates are missing at random. The approach facilitates the use of flexible machine-learning strategies for component nuisance functions while maintaining appropriate convergence rates for valid asymptotic inference. EHR data from Kaiser Permanente are used as a basis for extensive simulations that verify robustness properties in a wide range of realistic settings. The same data are also used to demonstrate the use of the method to analyze differences between two common bariatric surgical interventions for long term weight and glycemic outcomes among a cohort of severely obese patients with type II diabetes mellitus.

Additional Authors: Luke Benz, Rajarshi Mukherjee, Rui Wang

Session: Instrumental variables (Friday 11th April, 09:00-10:40)

Location: Auditorium 2

Abstract author: Christian Tien

Affiliation: QuantCo

Title: Identification and Estimation with Deconfounded Instruments under Index Sufficiency

Abstract: This paper extends a novel methodology, called common confounding (CC), for identifying and estimating the causal effects of endogenous (treatment) variables on an outcome variable with partially endogenous instrumental variables. A crucial estimation step called deconfounding recovers variation in the instruments, which is unassociated with some observed variables called proxies, and consequently with any unobserved variables that explain the association between the instruments and proxies. These unobserved variables are called common confounders of the instruments and proxies. This paper explores the role of index sufficiency as a minimal parametric assumption, which naturally fits in with the deconfounding approach and permits the identification and estimation of causal effects in the common confounding setup. Novel semiparametric estimation theory is provided. Root-n-estimation of causal effects under index sufficiency in otherwise nonparametric models is shown to be possible, combining ideas from debiasing with respect to sequentially dependent nuisance functions [Singh, 2021, Chernozhukov et al., 2022] with recent results on strong identification subject to nuisance functions, which are defined as solutions to possibly ill-posed inverse problems [Bennett et al., 2022]. An empirical application on the returns to education with NLS97 data demonstrates the appeal of this approach in practical settings with partially endogenous instruments.

Session: Instrumental variables (Friday 11th April, 09:00-10:40)

Location: Auditorium 2

Abstract author: Dean Knox

Affiliation: UPenn

Title: A causal ablation framework for evaluating assumptions

Abstract: We formalize an causal ablation framework for characterizing the role of various assumptions in the conclusions of an analysis. The ablation framework proceeds by systematically weakening individual assumptions on which the analysis relies, then utilizing partial identification techniques to evaluate how sharp bounds widen. The first step in the framework is to enumerate component assumptions in the baseline analysis—such as exclusion, monotonicity, and homogeneity in an instrumental variable study—along with a totally ordered "weakening sequence" for each assumption. Second, we construct a hierarchy of all assumption sets, which is partially ordered in strength: the baseline assumption set is stronger than both (i) an alternative that weakens assumption 1 and (ii) another that weakens assumption 2, but (i) and (ii) cannot be ranked because differing assumption types are non-comparable. Third, we organize this hierarchy using its Hasse diagram, a useful graphical depiction in which each "ablation path" represents one possible sequence through which assumptions can be iteratively pared down. We conduct falsification tests and estimate sharp bounds at each point, representing one possible assumption set. Finally, we identify the weakest assumption sets under which direction or magnitude of the effect can be determined, and we summarize the importance of component assumptions based on whether they are necessary, sufficient, or irrelevant for drawing each type of conclusion.

Additional Authors: Luke Keele, Ilya Shpitser

Session: Instrumental variables (Friday 11th April, 09:00-10:40)

Location: Auditorium 2

Abstract author: Jordan Penn

Affiliation: King's College London

Title: Partial identification with mostly invalid instruments and homogenous treatment effects

Abstract: Instrumental variables (IVs) are widely used to estimate treatment effects in the presence of unobserved confounding between exposure and outcome. An IV must affect the outcome exclusively through the exposure and be unconfounded with the outcome. I show in nonparametric structural causal models (SCMs) with homogenous treatment effects that a single latent statistic captures which candidate IVs satisfy these assumptions and summarizes "how invalid" each remaining candidate is. I discuss conditions for (partial) identification of the average treatment effect (ATE) when any combination of bounds regarding how many and/or which and/or to what degree candidate IVs may be invalid. In general, the resulting sharp identified region for the ATE is nonconvex and I discuss its geometry in the cases of continuous or discrete, bounded or unbounded outcome domains. Under finite sample uncertainty, I provide a means for forming asymptotically valid confidence sets for the ATE. I study the implications for inference in nonlinear and linear SCMs using simulated data and summary statistics previously used in a Mendelian randomization experiment.

Additional Authors: Lee Gunderson, Gecia Bravo-Hermsdorff, David S. Watson, Ricardo Silva

Session: Instrumental variables (Friday 11th April, 09:00-10:40)

Location: Auditorium 2

Abstract author: Andrej Srakar

Affiliation: Institute for Economic Research Ljubljana

Title: Instrumental Variable Estimation in Compositional Regression

Abstract: Time use surveys are used in many areas of economics, including economics of health and long-term care. If analyzed in a regression context, time use survey data suffer from the problem of spurious correlation noted in early works of Aitchison (1986). This problem leads to a need for compositional regression perspective on a geometric simplex. We develop an instrumental variable compositional regression model, building on two strands of literature with applications for health economics and economics of long-term care. We extend Florens and Van Bellegem (2015) functional instrumental variables model to compositional data setting where either or both independent and dependent variables are of compositional nature. We show there exist two ways of deriving compositional IV's, one using isometric log-ratio transform and Chesher et al. (2013)'s IV model of multiple discrete choice; and another deriving from the recent literature on compositional functional data in Bayes spaces (Machalova et al., 2021). We follow the latter and show that estimation, similar to the one of Florens and Van Bellegem leads to an ill-posed inverse problem with known but data-dependent operator. We resolve this in a context of multiplication by an instrument-dependent operator and by a penalized least squares estimation, and we also extend the notion of instrument strength to compositional setting. We establish appropriate functional central limit theorem in this context of Bayes spaces instead of more conventional Hilbert spaces. and study the finite sample performance in a Monte Carlo simulation setting. Our application studies relationship between long term care for older people and paid work, using recent time use survey from Survey of Health, Ageing and Retirement in Europe (SHARE).

Session: Instrumental variables (Friday 11th April, 09:00-10:40)

Location: Auditorium 2

Abstract author: Luca Locher

Affiliation: ETH Zurich

Title: Identity slippage: characterizing a systematic bias in applications of instrumental variables

Abstract: The local average treatment effect (LATE) is a popular parameter in applied domains, owing to its enhanced identifiability relative to the ATE. Understood this way, the LATE is part of a growing class of pragmatic causal estimands that are increasingly supported by methodological literatures in statistics, econometrics and other allied theoretical fields. This class includes interventional mediation parameters, overlap-weighted ATEs, and an expanding set of data-adaptive and projection-based analogues of classical causal parameters. These parameters are designed to avoid assumptions that are not often supported in applications, like positivity, cross-world independences, latent homogeneity of conditional effects, and non-exceptional laws, to name a few, thus offering practical advantages to investigators. A common critique, however, is that these estimands are challenging to interpret and thus may have diminished relevance to high-stakes policy decisions. One concern is that applied authors may systematically misinterpret pragmatic estimands, as if estimates corresponded to their classical analogues. Such errors may be of substantial consequence whenever pragmatic estimands and their analogues take values corresponding to qualitatively different policy actions. To test this hypothesis we conducted an analysis of interpretational errors in the applied instrumental variable (IV) literature. Borrowing methods typically used in meta-analysis, we targeted the population of articles applying IV methodology published between 2019 and 2023 in the medical, epidemiological, political and economic sciences (n=311 unique studies). A preregistered study protocol is available on OSF under the registration id 'fxq5v'. We find that a large majority of studies targeted the LATE, although specific interest in this parameter was rare. Almost two-thirds of these studies contained claims that mistakenly suggested that the marginal ATE was targeted. These phenomena were highly prevalent across disciplines. We also present re-analyses of several of the included studies; sharp-bounds on the ATE under the assumed models in these studies illustrate the potential practical consequences of these interpretational errors: in several cases, possible values of the partially-identified ATE qualitatively diverge from the LATE point estimate. Results suggest that the validity of conclusions drawn from results of IV applications is often compromised by interpretational errors. Our findings are consistent with an isomorphic analysis of applications investigating causal mediation that pragmatically targeted stochastic interventional parameters; therein, results were frequently interpreted as natural mediation parameters, which were not identified.

Additional Authors: Aaron Sarvet

Session: Invited (Friday 11th April, 11:10-11.40)

Location: Auditorium 1

Abstract author: Antoine Chambaz

Affiliation: Paris Cité University

Title: Physics-Informed Learning of a Plausible Explanation for Dysfunction

Abstract: TBA

Session: Keynote (Friday 11th April, 11:40-12:40)

Location: Auditorium 1

Abstract author: Sonja Swanson

Affiliation: University of Pittsburgh

Title: Methods matter

Abstract: When access to a highly lethal and commonly used suicide method is reduced in a population, that population's overall suicide rates decline. Estimating the extent to which specific interventions and policies might reduce suicide risk, however, generally relies on observational data. I will present work in this topic area, highlighting the need for careful attention to causal effect estimation in settings for which trials are not feasible and even observational data are especially limited. In particular, I will present work that builds upon the instrument-based partial identification literature (e.g., see review: Swanson et al. JASA 2018) with a novel extension applied to policy evaluation.