EuroCIM 2025 Poster Presentations

List of Abstracts

12 **Diversifying Clinical Trials with Adaptive Targeted Maximum Likelihood Estimation (A-TMLE): A Data Fusion Approach for Real-World Evidence** **Rachael Phillips**

13 **Causal Record Linkage: Critical Issues and Novel Approaches to False Discovery Propagation** **Kayané Robach**

14 **Causal Case-Mix: a novel framework for understanding prediction model performance under changes in case-mix** **Wouter van Amsterdam**

15 **Addressing Positivity Violations in Continuous Interventions through Data-Adaptive Strategies** **Han Bao**

16 **Optimal sequential decision-making with initiation regimes** **Julien Laurendeau**

17 **Causal Inference Under Stochastic Interventions Inducing Missingness-Specific Independencies** **Johan de Aguas**

18 **Causal mediation analysis with spatial interference** **Chiara Di Maria**

19 **Learning and Testing Exposure Mappings of Interference** **Mara Mattes**

20 **Bootstrapped Conformal Prediction on the Surrogate-index Framework to Quantify Prediction Uncertainty around Predicted Long-term Treatment Effects** **Tung Nguyen Huy**

21 **Interference in Competing Infectious Variants** **Gellert Perenyi**

22 **Estimating treatment effects in the presence of interference using domain adversarial training** **Daan Caljon**

23 **Method-of-Moments Inference for GLMs and Doubly Robust Functionals under Proportional Asymptotics** **Xingyu Chen**

24 **Prognostic scores and representation learning for causal effect estimation with weak overlap** **Oscar Clivio**

25 **Causal inference targeting a concentration index for studies of health inequalities** **Mohammad Ghasempour**

26 **Predicting Treatment Effects Across Contexts for Perturbational Biology** **Luka Kovačević**

27 **Learning to rank treatment effects** **Myrto Limnios**

28 **Performance of Cross-Validated Targeted Maximum Likelihood Estimation** **Miguel-Angel Luque Fernandez**

29 **Matching Methods for Difference-in-Differences with Multiple Time Periods: Evaluating the Equality of ATT Estimates Across Time** **Junho Jang**

30 **High-Dimensional Matching with Genetic Algorithms** **Hajoung Lee**

31 **Evaluating time-specific treatment effects using randomization inference** **Sangjin Lee**

32 **More Powerful Selective Inference in Propensity Score Analysis** **Sarah Pirenne**

33 **Selection bias due to omitting interactions from inverse probability weighting** **Liping Wen**

34 **Estimating Average Treatment Effect via Marginal Outcome Density Ratio** **Linying Yang**

35 **Causal discovery for multi-cohort studies** **Christine Bang**

36 **Nonparametric conditional independence testing with mixed data** **Luca Bergen**

37 **Scalable Constraint-Based Causal Discovery with Multiple Imputation for Incomplete Data** **Frederik Fabricius-Bjerre**

38 **When Causal Discovery Errs: Inconsistency Detection for PC-like Algorithms** **Sofia Faltenbacher**

39 **Small Open-Source LLMs for Causal Discovery [chatGPT-generated by local organizers as no title was provided]** **Louis Hernandez**

40 **Score-Based Causal Discovery with Temporal Background Knowledge** **Tobias Ellegaard Larsen**

41 **Full pipeline for causal inference on limited sample size** **Tobias Strømgren**

42 **Measuring similarity of causal graphs** **Jonas Wahl**

43 **Selection bias of cause-specific hazard ratios: the impact of competing events** **Mari Brathovde**

44 **Decision Analytical Models as Causal Models** **Maurice Korf**

45 **Discussion on treatment effect waning** **Eni Musta**

46 **A Diagnostic to Find and Help Combat Positivity Issues— with a Focus on Continuous Treatments and Modified Treatment Policies** **Katharina Ring**

47 **Post-treatment problems: What can we say about the effect of a treatment among sub-groups who (would) respond in some way?** **Tanvi Shinkre**

48 **Quantification of Vaccine Waning as a Challenge Effect** **Matias Janvin**

49 **Expressing Cost of Causal Assumptions Through Partial Identification** **Jakob Zeitler**

50 **Target trial emulation in biobank data: estimating the effect of cholesterol-lowering therapy and PRS in the Estonian Biobank** **Saskia Kuusk**

51 **A Graphical Approach to State Variable Selection in Off-policy Learning** **Joakim Blach Andersen**

52 **Theory and Application of longitudinal causal inference for comparative effictiveness in diabetes research** **Thomas Gerds**

53 **Model-free estimands for target trial analysis** **Edoardo Efrem Gervasoni**

54 **Structural Nested Models in Target Trial Emulation** **Fuyu Guo**

55 **Improving inference procedures in the R package TrialEmulation with resampling methods** **Juliette Limozin**

56 **Analysis of Longitudinal Causal Effects Using Joint Logit Models** **Ji Luo**

57 **A visual guide to g-estimation to handle symptomatic treatment at multiple timepoints in a randomized controlled trial** **Baldur Magnusson**

58 **Target trial emulation of dual treatment analyses: challenges of longitudinal TMLE** **Anna Menacher**

59 **Handling informative patient monitoring in routinely-collected data used to estimate treatment effects, with application to high-frequency hospital data** **Leah Pirondini**

60 **A comparison of statistical and computational tradeoffs in Sequential Target Trial Emulation approaches** **Carlos Poses**

61 **HOW DOES MODE OF BIRTH AFFECT THE RISK OF DEATH OR BRAIN INJURY IN VERY PRETERM BABIES? METHODOLOGICAL CONSIDERATIONS FOR A TARGET TRIAL EMULATION APPROACH** **Nicola Reeve**

62 Target Trial Emulation for the effect of Dialysis-duration on the Survival Benefit of Kidney Transplant Dries Reynders

63 Emulating non-inferiority target trials: challenges and considerations Sabine Landau

64 Improving interim decisions for single-arm trials by adjusting for baseline covariates and intermediate endpoints Eline Anslot

65 Unveiling Sleep Dysregulation in Chronic Fatigue Syndrome (CFS) with and without Fibromyalgia (FM) through Bayesian Networks Michal Bechny

66 kernel matching based on GPS in studying short-term effects of air pollution on population health Giulio Biscardi

67 Assessing Principal Causal Effects with Outcome-dependent Sampling Using Principal Score Methods: An Application to the E3N Cohort Lisa Braito

68 External reproduction of a proxy-based causal model for estimating average effects of sequential vs concurrent chemo-radiotherapy on survival in stage III Non-Small Cell Lung Cancer (NSCLC) Charlie Cunniffe

69 Integrating causal reasoning in academic curricula: exploring barriers and solutions Annick De Paepe

70 Inference on sustained treatment strategies, with a case study on young women with breast cancer Elise Dumas

71 Policy learning under constraint: seeking primary outcome maximization under adverse outcome control Laura Fuentes Vicente

72 Effect modification in studies affected by "truncation by death" Bronner Gonçalves

73 Refining the Allostatic Self-Efficacy Theory of Fatigue and Depression Using Causal Inference Alex Hess

74 The impact of dynamic and modified hemodynamic treatment policies on intraoperative hypotension: An application of causal inference methods in Anaesthesiology Markus Huber

75 Hybrid approaches, case studies in methods leveraging external data in randomized clinical trials Rima Izem

76 Effect of COVID-19 Vaccination Initiation on One-Year All-Cause Mortality Among Elderly Estonians: A Sequential Target Trial Emulation Anastassia Kolde

77 Treatment effect estimation with counterfactual prediction using individual treatment plans: theory and application in radiotherapy Lotta M. Meijerink

78 Use of Causal Inference in a Vaccine Development Program: A Case Study Joris Menten

79 An examination by Principal Component Analysis of the effect of timing of food intake on BMI in the INRAN-SCAI 2005/06 nutrition survey Luigi Palla

80 Thoughts after Droughts: Combined effects of prenatal and postnatal drought-related nutritional challenges on adolescent cognitive health in rural India Fabienne Pradella

81 Data-Driven Causal Discovery in Lifestyle Modelling: Insights from a Longitudinal Study with Wearable Data Radoslava Svihrova

82 Synthetic control via covariate balancing with applications in long-term health impacts of extreme climate events Xiao Wu

83 "Explaining" the gender pay gap Christiane Didden

84 Smoking heat Gun: estimating the effect of heatwaves on gun violence through a distance augmented synthetic control method Giulio Grossi

85 Assessing the Causal Impact of Early Tracking Postponement on Inequality of Opportunity: Evidence from the Italian Single Middle School Reform Kevin Taglialatela Scafati

86 A Nonparametric Bayesian Approach for High-Dimensional Causal Effect Estimation in Survival Analysis Tijn Jacobs

87 Population average treatment effect estimation under positivity violations: Adapting BART+SPL for high-dimensional covariates Lennard Massmann

88 Type 2 Tobit Sample Selection Models with Bayesian Additive Regression Trees Eoghan O'Neill

89 Bayesian estimation of heterogenous treatment effects from panel data Pantelis Samartsidis

90 Personalised Causal Effect Estimation with Observational and Experimental Data Sofia Triantafillou

91 Causal Inference on Quantiles in High Dimensions: A Bayesian Approach Duong Trinh

92 Causally Sound Priors for Binary Experiments Nicholas Irons

93 **Semiparametric Triple Difference Estimators** **Sina Akbari**

94 **Causal Effects of Time-Varying Exposures: A Comparison of Structural Equation Modeling and Marginal Structural Models in Cross-Lagged Panel Research** **Jeroen Mulder**

95 **The Impact of COVID-19 Pandemic on Non-deferrable Diseases** **Sara Muzzì**

96 **A joint test of unconfoundedness and common trends** **Eva Oeß**

97 **Conditional Independence Testing in Time Series** **Jieru Shi**

98 **Difference-in-Differences with Non-Ignorable Attrition** **Javier Viviens**

99 **Enhancing Generalizability in Regression Discontinuity Designs: A Graphical Approach to Combining Multiple Fuzzy RDDs** **Julia Kowalska**

100 **Differential Network inference based on causal discovery** **Emilie Devijver**

101 **Weak equivalence of causal graphs** **Søren Wengel Mogensen**

102 **The Case for Time in Causal DAGs** **Alexander Reisach**

103 **Revisiting Model Evaluation in Treatment Effect Estimation: A Critical Appraisal** **Hugo Gobato Souto**

104 **Shrinkage Bayesian Causal Forest with Instrumental Variable** **Jens Klenke**

105 **Targeting relative risk heterogeneity with causal forests** **Vik Shirvaikar**

106 **The Parachuted Hybrid CATE Estimator with Bootstrap Methods for Inference** **Xianlin Sun**

107 **Negative Control Outcomes for Selection Bias in Instrumental Variable Analysis** **Apostolos Gkatzionis**

108 **Estimating Heterogeneous Causal Effects with Tree-Based Methods under Imperfect Compliance and Positivity Violations** **Karolina Gliszczynska**

109 **Assessing changes in cell composition in observational single-cell gene expression studies** **Amber Huybrechts**

110 **Understanding spatial models with causal inference:** **Tim Lucas**

111 **E-value for RMTL: Evaluating Outcomes Under Non-Proportional Hazards** **Carmen Reep**

112 **Confounding of the competing event in time-to-event analyses** **Jost Viebrock**

113 **A promotion time cure model with endogeneity simultaneously affecting the cure probability and survival time of the susceptibles** **Lasse Winding**

114 **The role of estimation in Mendelian randomization: should Mendelian randomization investigations provide estimates?** **Benjamin Woolf**

115 **Interventional effects modeling in telework-health research: exploring the indirect and direct effects in the telework work-engagement relationship** **Eduardo A. Bracho Montes de Oca**

116 **Accessible methods for causal mediation analysis in two-level data structures** **Kirsty James**

117 **An adaptive test for natural indirect effect in large-dimensional mediation analysis** **Feng Liang**

118 **Longitudinal Mediation Analysis of Multiple Mediators: Estimating Brain Region-Specific Effects in Major Depressive Disorder** **Can Long**

119 **Mediation analysis in longitudinal intervention studies with an ordinal treatment-dependent confounder** **Mikko Valtanen**

120 **Performance of G-computation estimators in randomized controlled trials with high-dimensional covariates** **Muluneh Alene Addis**

121 **Estimating treatment effect in randomized trial after control to treatment crossover using external controls** **Christiana Drake**

122 **Powering RCTs for marginal effects with GLMs using prognostic score adjustment** **Emilie Højbjerre-Frandsen**

123 **Federated Causal Inference: Multi-Site ATE Estimation beyond Meta-Analysis** **Rémi Khellaf**

124 **Bayesian covariate-adjusted response-adaptive intervention allocation in stepped-wedge cluster randomised trials** **Constantin Schmidt**

125 **Establishing when to use causal machine learning for conditional average treatment effect estimation in randomised controlled trials using simulation** **Eleanor Van Vogt**

Abstract number: 1

Date: Wednesday 9th April

Title: Disentangling Anticipatory Effects and Intervention Impacts in Policy Evaluation Studies

Presenter: Damiano Baldaccini

Affiliation: University of Milano-Bicocca and University of Florence

Topic: Applications in Economics

Abstract: In the field of Policy Evaluation, the primary objective is to assess the effects of a policy or intervention on a specific outcome or set of outcomes.Under the Rubin Causal Model the effects of interests are defined as comparisons of the observed outcomes under the intervention with counterfactual outcomes, those that would have occurred in the absence of the intervention. A significant challenge in this context arises when the policy affects all units simultaneously, thus eliminating the possibility of a valid control group.Recent advancements in counterfactual evaluation have addressed this issue by developing methods for estimating the impact of interventions or shocks over time, without relying on traditional control groups. The central idea behind these methods is to leverage pre-intervention data to predict a hypothetical counterfactual scenario that represents the absence of the intervention. The predicted scenario is then compared with the observed outcomes post-intervention to estimate the policy's effects.Anticipatory effects refer to behavioural or systemic changes that occur prior to the actual implementation of a policy, in response to its announcement. Understanding the evolution of these effects over time is often critical when evaluating public policies.We formally define anticipatory effects and the effects of the intervention and introduce a set of assumptions to identify and disentangle anticipatory effects. We evaluate the performance of the proposed methodology through a series of simulations under different scenarios about anticipatory effects and intervention effects.We apply the methodology to an empirical analysis of vehicle purchase data in Italy, focusing on the impact of a policy that introduced a dual incentive-disincentive mechanism referred to as a feebate system. This system offers financial incentives for the purchase of low-emission vehicles while imposing disincentives, such as taxes, on high-emission vehicles. This application underscores the practical utility of our approach in understanding the temporal dynamics of the policy's effects on consumer behaviour.

Additional Authors: Alessandra Mattei (University of Florence), Fiammetta Menchetti (University of Florence), Patrizia Lattarulo (IRPET)

Abstract number: 2

Date: Wednesday 9th April

Title: Should I Stay or Should I Go? Exploring Family and Employment Influences on Young Adults Leaving Home

Presenter: Ibone Bilbao

Affiliation: University of the Basque Country

Topic: Applications in Economics

Abstract: In Spain, young individuals delay leaving their parental households until their 30s, placing the country among those in Europe with the latest average age of emancipation. This article explores how various labor market transitions and familial shocks affect the timing of nest-leaving. Using duration analysis and causal inference techniques, we find that labor transitions—such as moving from inactivity to employment, from temporary to permanent contracts, and from part-time to full-time work—reduce the average emancipation age by 4.54 years, 3.92 years, and 5.67 years, respectively. Moreover, paternal transitions from inactivity or unemployment to employment significantly decrease children's emancipation age by 5.83 years, whereas maternal employment transitions have no significant effect. Conversely, maternal health shocks delay emancipation by 1.15 years, while paternal health shocks show no significant impact.

Additional Authors: Javier Gardeazabal and Alaitz Artabe

Abstract number: 3

Date: Wednesday 9th April

Title: Who benefits from tax windfalls? Incidence and rent sharing in small firms

Presenter: Haotian Deng

Affiliation: Department of Economics, Ghent University

Topic: Applications in Economics

Abstract: This paper empirically synthesizes payroll tax incidence and rent sharing in small firms in a unified framework. We exploit an unanticipated payroll tax cut (subsidy) reform in Belgium targeting entrant employer firms that just hired their first employees. This reform generated a temporary tax windfall gain disposable at the firm level. We employ triple differences as our identification strategy to document dynamic effects of this subsidy reform on eligible entrants over their employer duration. At the employee level, the reform reduced firms' labor costs by 200–400 euros per full-time equivalence (FTE) per quarter during the temporary subsidy periods, leaving average wage rates unaffected. Firm-level employment expanded gradually over the subsidized periods and grew by 0.1 FTE (6.7%) after the subsidy expired, while their average economic rents did not change. Our preliminary findings contradict the conventional wisdom in the literature of wage incidence and rent sharing but we argue that the average null effects are masked by essential heterogeneity. We plan on two next steps to enrich our study. First, we will empirically test whether our prediction on firms' heterogeneous responses is true or not: low-productivity firms mainly used the subsidy to secure employment while high-productivity ones afforded to generate higher profits and pay higher wages. Second, we will apply an instrumental variable approach that directly follows from the triple differences strategy with the Wald estimator, and identify three important economic parameters among entrant employer firms: the subsidy-employment elasticity, the subsidy-wage pass-through, and the rent-sharing elasticity.

Additional Authors: Sam Desiere (Ghent University, IZA); Bart Cockx (Ghent University, IZA, Catholic University of Louvain, CESIfo, Maastricht University)

Abstract number: 4

Date: Wednesday 9th April

Title: Wind of Structural Change - The Economic Integration of East Germany and Occupational Status Gaps

Presenter: Raffael Kind

Affiliation: European University Institute

Topic: Applications in Economics

Abstract: The reunification of Germany in 1990 triggered a decade of profound structural transformation in East Germany. The transition from a socialist planned economy to a capitalist market economy led to rising wages but also job losses and downward occupational mobility, as many workers were forced into lower-status jobs due to firm liquidations and privatizations. To oversee the privatization of state-owned East German enterprises, the German government established the so-called Treuhand agency. It became the world's largest holding company overnight, managing approximately 4 million jobs — 2.5 million of which were lost during its existence between 1990 and 1994. Further, the Treuhand became infamous for preventable firm liquidations. This study examines how the rapid economic transformation following reunification affected East Germans' occupational status. For this purpose, exploiting geographic variation in the intensity of Treuhand liquidations, we apply the method proposed by Borusyak & Hull (2023). This approach is well-suited to our context, as counties were non-randomly exposed to the exogenous firm closure shocks due to factors such as counties' sectoral composition and economic geography. The non-random exposure invalidates two conventional identification strategies. First, instrumenting the exposure of Treuhand liquidations with the firm closure shocks fails to account for the non-random nature of counties' exposure. Second, controlling for the factors governing this non-randomness (sectoral composition, location, etc.) cannot fully capture the complex spatial dependence of the exposure to Treuhand liquidations on these factors. In contrast to Borusyak & Hull (2023), our identification does not rely on quasi-randomness in the timing of the exogenous shocks but on quasi-randomness in their intensity across counties. Our findings suggest that the economic transformation following reunification set many East Germans on a path of occupational status decline. This adds a new perspective on the effects of macroeconomic transformations. Examining individual labor market experiences through the lens of occupational status provides valuable insights into the social fears and resentments that have shaped post-reunification dynamics in East Germany. These patterns also resonate with similar dynamics in other Western countries undergoing economic upheavals, where the German reunification's unique context cannot serve as an identification strategy.

Additional Authors: Robert Schall (EUI), David Wittekopf (EUI)

Abstract number: 5

Date: Wednesday 9th April

Title: Gender Differences in Healthcare Utilization: Evidence from Unexpected Adverse Health Shocks

Presenter: Nadja van 't Hoff

Affiliation: University of Southern Denmark

Topic: Applications in Economics

Abstract: This paper is the first to causally quantify gender differences in healthcare utilization to better understand the male-female health-survival paradox, where women live longer but experience worse health outcomes. Using rich Danish administrative healthcare data, we apply a staggered difference-in-differences approach that exploits the randomness in treatment timing to estimate the causal impact of adverse health shocks, such as non-fatal heart attacks or strokes, on healthcare use. Our findings suggest that men consistently use more healthcare than women, shedding light on the underlying factors driving gender disparities in health outcomes. These insights contribute to the broader discourse on healthcare equity and inform policy interventions aimed at addressing these imbalances.

Additional Authors: Giovanni Mellace, Seetha Menon

Abstract number: 6

Date: Wednesday 9th April

Title: Beyond Social Mobility - The Gender-Class Pay Gap

Presenter: David Wittekopf

Affiliation: European University Institute, Florence, Italy

Topic: Applications in Economics

Abstract: We investigate whether socioeconomic family background differentially impacts labor market outcomes for women and men. Specifically, we define and estimate a pay disparity we refer to as the Gender-Class Pay Gap. To this end, we compare established decomposition methods (Oaxaca (1973); Gelbach (2016)) with counterfactual mediation analysis (e.g., Blakely et al. (2018)) to decompose this gap. We argue that mediation analysis offers a more credible approach to gap decomposition than conventional methods, as it accounts for the underlying temporal structure. Traditional decomposition methods treat controls as if they occur before or simultaneously with the treatment. In contrast, mediation analysis allows for a distinction between pre-treatment controls (referred to as controls) and post-treatment "controls" (referred to as mediators), providing a clearer understanding of causal pathways. This distinction enhances the credibility of the analysis and sheds light on the mechanisms driving our findings. Further, our approach allows for including controls in the potential outcome framework. Using data from the German Socioeconomic Panel (2010–2019), our results reveal that even after achieving upward social mobility, individuals from lower socioeconomic backgrounds earn significantly less labor income than peers of comparable current socioeconomic status but higher socioeconomic family background. While the gap is alleviated when accounting for, among other factors, occupational and educational sorting, it remains both statistically and economically significant. Crucially, when analyzing women and men separately, we find that the class pay gap is almost entirely driven by men. Among women, the socioeconomic family background is not associated with significant differences in pay. Remarkably, however, women from high-class backgrounds earn significantly less than men from low-class backgrounds, even when controlling for occupational sorting, education, part-time work, and other relevant factors. This highlights a striking reality: even the most privileged women earn less than the least privileged men. We believe our findings offer a novel perspective on the interplay between gender and socioeconomic disparities in labor income, adding a new dimension to the discussion of the gender pay gap.

Additional Authors: Raffael Kind

Abstract number: 7

Date: Wednesday 9th April

Title: How to transport causal effects? A step-by-step approach derived from applied work

Presenter: Vincent Brugger

Affiliation: Center for Preventive Medicine and Digital

Topic: Data fusion/generalizability/transportability

Abstract: Transportability methods are a promising approach to using study results for inferences in distinct populations. However, current applications of transportability methods differ substantially in terms of approaches used to transport causal effects from one population to another. This heterogeneity limits the comparability of study results and makes it hard for subsequent applications to apply transportability methods as there seems to exist little guidance on which methods to use and which nuances to consider in a transportability analysis. To overcome these issues, we derive a list of common elements from applications of transportability identified in a previously conducted literature review that are addressed in applied research using transportability methods and discuss approaches to address each element. As a result, the presented step-by-step procedure provides guidance for researchers and contains related code in R to conduct transportability analyses. We illustrate the step-by-step procedure in an application example of a multicenter randomized trial on the treatment of Indomethacin to avoid post endoscopic retrograde cholangiopancreatography pancreatitis. This should help future users of transportability methods to arrive at coherent results, comply with the approaches previous authors used, and help them to reason their own analysis decisions
.

Additional Authors: Fabian Manke-Reimers, Til Bärnighausen, Stefan Kohler

Abstract number: 8

Date: Wednesday 9th April

Title: Exploring Hierarchical Causal Models to Address Subunit-Level Confounding in Generalizing Clinical Trial Results

Presenter: Evangelos Dimitriou

Affiliation: University College London

Topic: Data fusion/generalizability/transportability

Abstract: Hierarchical data arise in many applications where subunits (e.g., patients) are nested within units (e.g., hospitals). In such settings, unit-level variables (e.g., site-specific protocols) can influence subunit-level variables (e.g., individual patient outcomes), and vice versa. Hierarchical causal models (HCMs) have been developed to address causal questions in these complex data structures, providing tools to identify causal effects even in the presence of unobserved unit-level confounders. However, current HCM approaches do not address subunit-level confounding, which occurs when unmeasured subunit-level variables affect subunit-level treatment and outcomes within units.This limitation is particularly problematic when combining data from multiple sources, such as multi-site randomized controlled trials (RCTs) and observational studies. While RCTs offer strong internal validity by eliminating confounding through randomization, their external validity can be threatened due to participants being unrepresentative of the target population. Observational studies complement RCTs by providing broader population coverage but are prone to unmeasured confounding. When integrating these data sources to generalize trial results, subunit-level confounding—present in observational data—threatens causal validity.In this work, we investigate the implications of subunit-level confounding in hierarchical causal models using Bayesian nonparametric methods. Through simulation studies, we examine how varying the degree of confounding affects causal findings, highlighting the challenges it poses in integrating multi-source data. Our results emphasize the critical need for methods that explicitly account for subunit-level confounding to ensure valid causal conclusions, particularly when combining data from randomized trials and observational studies.

Additional Authors: Edwin Fong, Karla Diaz-Ordaz, Brieuc Lehmann

Abstract number: 9

Date: Wednesday 9th April

Title: ManyData: Combining multiple experimental or observational datasets through a power-likelihood

Presenter: Xi Lin

Affiliation: University of Oxford

Topic: Data fusion/generalizability/transportability

Abstract: Randomised controlled trials (RCTs), the gold standard for causal inference, often lack statistical power to examine effects in subpopulations or secondary outcomes. This has prompted growing interest in data fusion methods that integrate RCTs with observational data. Much of this research has focussed on combining an RCT with one other study, rather than more general approaches.In this work, we propose a novel method for combining a privileged data source with multiple secondary datasets to enhance the efficiency of treatment effect estimation. Our approach generalizes the existing power-likelihood method for data fusion. It is data-adaptive by optimizing the expected log predictive density (ELPD) to determine learning rates that effectively regulate information from secondary sources. We validate our method through extensive simulation studies, demonstrating reduced error in treatment effect estimates. Additionally, we apply it to real-world data, showcasing its practical utility.

Additional Authors: Xi Lin, Alexander Gruen, Robin Evans

Abstract number: 10

Date: Wednesday 9th April

Title: External validity of the community-based OPIC trials – exploring potential selection bias due to loss to follow-up and effect heterogeneity between trial arms

Presenter: Fabian Manke-Reimers

Affiliation: Center for Preventive Medicine and Digital Health, Medical Faculty Mannheim, Heidelberg University

Topic: Data fusion/generalizability/transportability

Abstract: IntroductionThe OPIC (Pacific Obesity Prevention in Communities) project consisted of obesity interventions in school-based cluster randomized trials in four countries (Australia, Fiji, New Zealand and Tonga). ProblemDue to high loss to follow-up (e.g., Fiji 41% follow-up rate) and shifted covariate distributions between censored and uncensored participants, the validity of intervention effects might suffer from selection bias. We aimed to investigate 1) whether the published results of a positive intervention effect on BMI z-score in Australia and negative intervention effect on quality of life in Fiji applies to the full baseline samples, and 2) the extent to which measured covariates can explain effect heterogeneity between the four trial countries.MethodsEffect sizes from the longitudinal subset of the data were generalized to the full baseline samples. Transportability analysis was conducted by transporting intervention effects from the participants with longitudinal data to the participants lost to follow-up and from one trial country to another. Three augmented inverse weighting estimators robust against model misspecification of the outcome or sampling model were applied for the risk difference in generalizability and transportability analyses, respectively. Bootstrap intervals were estimated with clustered standard errors. Applied covariate sets were 1) chosen based on a systematic data-driven approach and 2) covariate sets from the original analyses. Main resultsIn Australia, the average treatment effect on BMI z-score remained significantly positive when applied to the whole baseline sample and the subset of participants lost to follow-up for both covariate sets. In contrast, negative intervention effects on quality of life were no longer apparent in the Fiji dataset, for one of three doubly estimators for the full baseline sample and for all three estimators in the subset of participants lost to follow-up, applying the new covariate set. Compared to the margin between point estimates in Australia and Fiji for the intervention effect on BMI z-score, applied doubly robust estimators yield point estimates within a 20% range of the targeted effect. ImplicationsOur results increase confidence in the positive intervention effect observed in Australia, and raise questions about the negative intervention effect on quality of life in Fiji.

Additional Authors: Vincent Brugger, Jane Jacobs, Vicki Brown, Steven Allender, Melanie Nicols

Abstract number: 11

Date: Wednesday 9th April

Title: Benchmarking observational analysis against registry-based trials: estimands, assumptions and estimators

Presenter: Camila Olarte Parra

Affiliation: Karolinska Institutet

Topic: Data fusion/generalizability/transportability

Abstract: To increase our confidence in observational analyses, we can compare real trials with observational emulations of target trials designed to be similar. Obtaining equivalent results could make observational analysis more trustworthy to evaluate other (long-term) outcomes outside the scope of the real trial and to assess treatment heterogeneity. Here, we outline a framework for formal benchmarking in an ideal setting of a registry-based trial, where the real trial is embedded in a nationwide registry. We describe the conditions required for valid comparisons and propose methods to evaluate the implications of these conditions on the observed data. We evaluate the proposed methods through plasmode simulations which allow us to generate datasets that mimic the features of real-world observational data while embedding known causal effects as found in trials. These simulations are designed to resemble a large cardiovascular trial nested in a large Swedish registry, where we have follow-up data for patients who did not enrol in the trial. Well-performing methods in simulations are subsequently applied to the real data from the Swedish registry, exploiting the rich variables available from the linkage of different population registers that contain clinical and socio-economic variables.

Additional Authors: Anita Berglund, Issa Dahabreh

Abstract number: 12

Date: Wednesday 9th April

Title: Diversifying Clinical Trials with Adaptive Targeted Maximum Likelihood Estimation (A-TMLE): A Data Fusion Approach for Real-World Evidence

Presenter: Rachael Phillips

Affiliation: University of California, Berkeley (UCB)

Topic: Data fusion/generalizability/transportability

Abstract: Increasing the diversity of clinical trial populations is critical to improving the generalizability of findings and addressing under-representation in medical research. However, pooling data from multiple sources to achieve this goal presents challenges, particularly in maintaining statistical efficiency while avoiding bias. This work introduces and explores the application of Adaptive Targeted Maximum Likelihood Estimation (A-TMLE) as a solution to these challenges, enabling efficient and unbiased data fusion for trial diversification in regulatory settings. We anchor our investigation on a high-impact clinical example: diversifying cardiovascular outcome trials (CVOTs) at Novo Nordisk to include historically underrepresented subgroups. Specifically, A-TMLE is used in real-world evidence studies integrating randomized control trial (RCT) data (LEADER, NCT01179048; DEVOTE, NCT01959529), and real-world data (Optum), accounting for differences in data quality, missingness, and population overlap. We also consider the data fusion application within an RCT, viewing all data outside the subgroup as external data. Through simulations and real-world analyses, we show how A-TMLE yields an equally valid estimator as the subgroup-specific estimator and we demonstrate its robustness for controlling Type I error while gaining efficiency by using external data without inducing bias. Our findings highlight the potential of this method to operationalize trial designs that are both scientifically rigorous and equitable, providing critical insights for regulatory science and clinical trial methodologies. This work emphasizes the importance of leveraging advanced causal inference tools to support evidence generation for diverse populations. By addressing the dual imperatives of efficiency and inclusivity, A-TMLE offers a scalable framework for meeting regulatory requirements while advancing health equity in clinical research.

Additional Authors: Sky Qui (UCB), Nerissa Nance (Novo Nordisk), Maya Petersen (UCB & UCSF) Mark van der Laan (UCB)

Abstract number: 13

Date: Wednesday 9th April

Title: Causal Record Linkage: Critical Issues and Novel Approaches to False Discovery Propagation

Presenter: Kayané Robach

Affiliation: Amsterdam UMC

Topic: Data fusion/generalizability/transportability

Abstract: From the idea of assembling a book of life for individuals to support public health decisions, record linkage emerged. It evolved to empower research by integrating information from diverse sources, such as electronic patient records and observational studies. Today, it offers an opportunity to control for confounding variables in causal inference, to study long term outcomes by linking longitudinal data, or else to explore secondary outcomes which were not measured in the first place.In this investigation, we review the perspectives taken on causal record linkage. Specifically, we examine uncertainty propagation from the linkage to the treatment effect estimation and we survey the two stage and joint approaches which are being used in applications to economics, banking, medicine. We give an account on the criteria addressed in recent developments and we evaluate the potential of a novel algorithm to control the rate of false links for causal effect estimation. Finally, we discuss the remaining challenges and outline potential directions for future causal research using linked data.

Additional Authors: Mark A. van de Wiel, Michel H. Hof, Stéphanie L. van der Pas

Abstract number: 14

Date: Wednesday 9th April

Title: Causal Case-Mix: a novel framework for understanding prediction model performance under changes in case-mix

Presenter: Wouter van Amsterdam

Affiliation: Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht

Topic: Data fusion/generalizability/transportability

Abstract: Prediction models inform important decisions in many domains. For example in healthcare, prediction models aid in diagnosis, prognosis, and treatment planning. When used in practice, these models require dependable performance, which for binary outcomes is typically assessed through discrimination and calibration. However, changes in the distribution of the data impact model performance and there may be important changes in the distribution between the last performance evaluation and the data a model is deployed on. In health-care, a typical change in distribution is a shift in case-mix. For example, for cardiovascular risk management, a general practitioner sees a different mix of patients than a cardiologist in a university hospital.This work introduces a novel framework that differentiates the effects of shifts in case-mix on a prediction model's discrimination and calibration based on the causal direction of the prediction task. When prediction is in the causal direction (often the case for prognosis predictions), calibration remains stable under case-mix shifts, while discrimination does not.Conversely, when predicting in the anti-causal direction (often with diagnosis predictions), discrimination remains stable, but calibration does not. These main results are illustrated with a simulation study and the framework's predictions are evaluated empirically on a dataset of 2030+ validations of 1300+ cardiovascular disease prediction models.The Causal Case-Mix framework offers guidance for the evaluation and deployment of prediction models in varying case-mix scenarios. By understanding the causal direction of the prediction task, stakeholders can anticipate changes in model performance metrics such as discrimination and calibration. This insight helps making decisions about model updates and ensuring reliable performance across different settings. Also, depending on whether the model's utility is more dependent on discrimination or calibration, prediction model developers my incorporate only causal or anti-causal features in their models. Ultimately, this framework aids in the development of robust prediction models.

Abstract number: 15

Date: Wednesday 9th April

Title: Addressing Positivity Violations in Continuous Interventions through Data-Adaptive Strategies

Presenter: Han Bao

Affiliation: Ludwig-Maximilians-Universität München

Topic: Dynamic treatment regimens

Abstract: Positivity violations pose a key challenge in the estimation of causal effects, particularly for continuous interventions. Despite their importance, existing methodologies provide limited tools for diagnosing and addressing these issues. Current approaches to handling positivity violations involve projection functions or modified treatment policies that redefine estimands. While effective in many contexts, these methods can result in unnatural estimands, thereby leading to compromises in interpretability or unintended deviations from the original research question. To address this gap, we propose a practical framework that both diagnoses positivity violations and resolves them through a data-adaptive strategy that maintains the interpretability of the new estimands and alignment with the original research question. First, we introduce the non-overlap ratio, a novel diagnostic tool for detecting positivity violations in the context of continuous interventions. This metric is a function of the intervention which quantifies the proportion of units outside the high-density region of the conditional intervention distribution, providing a clear measure of the feasibility of the intervention. The non-overlap ratio enables researchers to gauge the extent of positivity violations across different intervention regions. Second, we propose a data-adaptive intervention strategy, termed the "most feasible" intervention strategy, to construct a new estimand designed to address positivity violations. This unit-based strategy operates as follows: for a given intervention of interest, we first assess whether the intervention value is feasible for each unit. For units with sufficient support, we adhere to the intervention of interest. However, for units lacking sufficient support—as identified through the assessment of the non-overlap ratio—we reassign each infeasible intervention to the nearest feasible value within the practical support. The goal of this approach is to mitigate positivity violations while preserving interpretability through minimal adjustments to the infeasible interventions. Through simulations, we demonstrate that our method effectively reduces bias in the presence of positivity violations, regardless of whether the outcome function is correctly specified or misspecified during the estimation process. Additionally, the method successfully recovers the standard estimand when positivity violations are absent. We further validate its practical utility using real-world data from the CHAPAS-3 trial, which enrolled HIV-positive children in Zambia and Uganda.

Additional Authors: Michael Schomaker

Abstract number: 16

Date: Wednesday 9th April

Title: Optimal sequential decision-making with initiation regimes

Presenter: Julien Laurendeau

Affiliation: EPFL

Topic: Dynamic treatment regimens

Abstract: Consider an optimal dynamic treatment regime, correctly identified from a perfectly executed sequentially randomized experiment. Even when the experimental results are generalizable to a future target population, there is no guarantee that the optimal regime outperforms human decision-makers; human experts can do better than the optimal regime whenever they have access to relevant information beyond the covariates recorded in the experiment. Motivated by this fact, we derive results on a new class of regimes called initiation regimes. These regimes follow human decision-makers until it is more beneficial to initiate a sequential optimal regime from that point onward, and are guaranteed to outperform both entirely human and entirely algorithmic decision-makers, e.g., based on reinforcement learning algorithms. Furthermore, we present modified experimental designs that identify the best initiation regimes, show how the best initiation regime can be identified from classical observational data with commonly invoked assumptions, and give estimation and statistical inference methodology for these regimes. To illustrate the practical utility of our methods, we consider initiation regimes in a case study on back pain treatment.

Additional Authors: Aaron L. Sarvet and Mats J. Stensrud

Abstract number: 17

Date: Wednesday 9th April

Title: Causal Inference Under Stochastic Interventions Inducing Missingness-Specific Independencies

Presenter: Johan de Aguas

Affiliation: University of Oslo

Topic: Dynamic treatment regimens

Abstract: The recovery of causal effects in structural models with missing data often relies on m-graphs, which assume that missingness indicators do not directly influence substantive variables. However, in many real-world settings, missing data can alter decision-making processes, as the absence of key information affects downstream decisions and states. To overcome this limitation, we introduce lm-SCMs and lm-graphs, which extend m-graphs by integrating a label set that represents relevant context-specific independencies (CSI), accounting for mechanism shifts induced by missingness. We define two causal effects within these systems: the Full Average Treatment Effect (FATE), which represents the effect had no data been missing, and the Natural Average Treatment Effect (NATE), which captures the exposure effect alone. We propose recovery criteria for these queries and present doubly-robust estimators for a simple graphical model inspired by a real-world application. Simulations highlight key differences between these estimands and estimators. Findings from the application case suggest a small effect of ADHD treatment upon academic achievement among Norwegian schoolchildren, with a slight shift in effect due to missing pre-tests scores.

Abstract number: 18

Date: Wednesday 9th April

Title: Causal mediation analysis with spatial interference

Presenter: Chiara Di Maria

Affiliation: University of Palermo

Topic: Interference

Abstract: Drawing causal conclusions from the analysis of observational data requires specific, sometimes untestable assumptions. One of the most common is the so-called Stable Unit Treatment Value Assumption (SUTVA), which requires that the response variable of a subject depends only on their own treatment assignment, not on those of other units, or, in other words, there is no interference between units. This assumption is unlikely to hold in many real-world settings, especially in the context of spatial data, where the impact of a subject's variable on other subjects is called spillover effect. Indeed, the treatment assigned to some units may also affect units which did not receive it: for example, in environmental studies, the pollution levels in one area may affect the health outcomes of individuals in adjacent areas, or in the case of environmental policies, regions close to the one where the policy is actually implemented may benefit from it as well. Causal inference in spatial settings has recently received the attention of some scholars, who proposed new estimands to address the complexities connected to spatial data. However, to the best of our knowledge, no one has considered the issue of mediation with spatial interference. In this work, we propose estimands for the direct and the indirect effects, taking into account the distance between units. We discuss the assumptions necessary for the identification of such effects and show the impact of ignoring spatial patterns on estimates. Finally, we provide an application to real data.

Additional Authors: Giada Adelfio

Abstract number: 19

Date: Wednesday 9th April

Title: Learning and Testing Exposure Mappings of Interference

Presenter: Mara Mattes

Affiliation: Heinrich Heine University Duesseldorf, Duesseldorf Institute for Competition Economics

Topic: Interference

Abstract: Interference or spillover effects occur when an individual's outcome (e.g., health) is influenced not only by their own treatment (e.g., vaccination) but also by the treatment of others. These effects pose challenges for evaluating treatment effects, particularly when arbitrary forms of interference are considered. Exposure mappings provide a structured framework to study interference effects in treatment evaluation by explicitly modeling the mechanisms through which the treatment statuses of social contacts within an individual's network influence their outcome. This permits separating interference effects from a treatment's direct effect, if one appropriately controls for differences in the probabilities of specific exposure mappings across individuals, for instance, due to different network structures. Researchers typically impose a priori exposure mappings of limited complexity, which may be insufficient or inaccurate in capturing all interference effects. In the first part of our study, we use graph neural networks (GNNs) to learn exposure mappings in a data-driven way which allows to capture all interference effects more accurately. GNNs are particularly well suited to this task because they can exploit the graph-structured data to learn dependencies and relations within a network. The exposure mappings are derived from the embeddings learned by the GNN.In the second part of this study, we introduce a data-driven, machine learning-based method to validate the learned exposure mapping and thus test the identification of the direct effect. If this mapping is valid such that any interference operates through them, then the treatment of the other individuals should be statistically independent of an individual's own outcome, conditional on the learned exposure mapping and the underlying network. This testable conditional independence relies on two key assumptions: (1) the exogeneity of treatment assignment and (2) the premise that the influence of other individuals' treatments on an individual's own outcome occurs solely through the exposure mapping.Finally, we investigate the performance of both parts, the estimation of the exposure mapping and the identification test, in a simulation study and provide an empirical application.

Additional Authors: Martin Huber, Jannis Kueck

Abstract number: 20

Date: Wednesday 9th April

Title: Bootstrapped Conformal Prediction on the Surrogate-index Framework to Quantify Prediction Uncertainty around Predicted Long-term Treatment Effects

Presenter: Tung Nguyen Huy

Affiliation: Zalando SE, Berlin

Topic: Interference

Abstract: n many policy-relevant evaluation scenarios the long-term, primary outcome is often difficult and costly to measure. Athey et al. (2019) address this problem by developing the surrogate-index framework that produces a predicted long-term effect estimate by combining the short-term surrogate metrics with pre-experiment information to predict the treatment effect on the long-term, primary outcome. They show that under certain assumptions the effect on the predicted long-term outcome is an unbiased estimate of the true long-term treatment effect. However, a key challenge remaining is how to assess the uncertainty around the predicted treatment effects especially in applied cases where the validity of the underlying assumptions are not guaranteed. While we can get an estimate on the predicted long- term effect, decision makers would like to understand the degree of uncertainty in the prediction in addition to the difference across experimental variants. To address this open challenge, we complement the surrogate-index framework with a model-agnostic approach to quantify prediction uncertainty. Our proposed method generating prediction intervals for the average treatment effect is inspired by conformal prediction and bootstrapping techniques. Having a single ground-truth observation of the long-term average treatment effect, we draw multiple bootstrap samples of the training data to approximate the distribution of absolute prediction errors. Akin to confidence intervals, we then take the desired percentile of prediction errors from this distribution and use this value to build the prediction intervals around the predicted effects. Our simulation validates that the true ATE is contained within the generated prediction intervals at a reasonable frequency even when the validity of the underlying assumptions is not guaranteed.

Additional Authors: Mats Maiwald, Patrick Doupe, Vladimir Fux, Zoltan Puha, Luca Scotton, Konstantin Basmer, Tomi Kasurinen, Antti Saarilahti, Mustafa Khandwawala, Omar Tawfik, Benjamin Tanz

Abstract number: 21

Date: Wednesday 9th April

Title: Interference in Competing Infectious Variants

Presenter: Gellert Perenyi

Affiliation: EPFL

Topic: Interference

Abstract: Outcomes of practical interest are often multivariate. For example, pathogens often exist in heterogeneous variants, such as subtypes and strains. It is essential to quantify treatment effects on these variants to guide prevention policies and treatment development. However, a key challenge is ensuring that variant-specific treatment effects are well-defined and interpretable in diverse settings, particularly in the context of infectious diseases where interference can occur.To address this challenge, we formalize causal estimands that quantify variant-specific effects. These formalizations clarify the interpretation of existing methods and reveal how they relate to policy-relevant effects. As an example, individuals are routinely assumed to be independent and identically distributed (iid) in randomized trials, even in infectious disease settings. Even though this assumption can be reasonable when the trial participants constitute a small subset of the total population, however it may not be generalized to population-wide interventions, such as vaccination programs. In such settings, additional dynamics like indirect protection through herd immunity arise, raising concerns about the external validity of trial-based effect estimates.To ensure the applicability of vaccine efficacy parameters in settings with interference, we derive explicit conditions that are both scientifically and practically justifiable for infectious disease contexts. These conditions ensure that commonly reported efficacy parameters correspond to well-defined causal effects, even in the presence of interference. Furthermore, our results offer alternative justifications for reporting estimands on the relative, rather than absolute, scale.We illustrate our findings by analyzing a large HIV1 vaccine trial, where there is interest in distinguishing vaccine effects on viruses with different genome sequences.

Additional Authors: Mats Julius Stensrud

Abstract number: 22

Date: Wednesday 9th April

Title: Estimating treatment effects in the presence of interference using domain adversarial training

Presenter: Daan Caljon

Affiliation: KU Leuven

Topic: Machine Learning in Causal Inference

Abstract: Individualized treatment effect estimation enables data-driven optimization of decision-making. Traditionally, different units are assumed to be independent, meaning they do not influence one another. However, in many real-world scenarios, spillover, or network, effects are present. For example, a vaccine not only directly benefits its recipient by reducing their risk of severe illness but also benefits people in their social circle (e.g., friends and family) by reducing transmission risks.Recent advancements in causal machine learning for spillover effect estimation often rely on predefined exposure mapping, which specifies how treatments of connected instances influence the outcome of a given instance. However, this assumption is generally unrealistic in real-world scenarios where the exact mechanisms of spillover effects are unknown. Therefore, we propose a novel causal machine learning approach that integrates Graph Neural Networks (GNNs) with domain adversarial learning to estimate heterogeneous treatment effects in the presence of interference. The representational power of GNNs enables the model to learn various exposure mappings from data, while domain adversarial learning is used to learn balanced (treatment-invariant) representations. This balancing effectively reduces the treatment assignment bias often present in observational data. In this work, we empirically examine the importance of using these powerful GNN layers and balanced representation in improving the accuracy of treatment effect estimation. To this end, we perform an extensive analysis on (semi-)synthetic data with varying levels of treatment assignment strength, homophily, spillover effect strength, and spillover effect complexity (e.g., heterogeneous vs. homogeneous effects). The results of this analysis indicate that homophily and the treatment assignment mechanism interact to create clusters of treated and untreated nodes within the network. This poses a significant challenge for graph learning methods that do not address treatment assignment bias, leading to inaccurate treatment effect estimates. However, we empirically show that our proposed causal machine learning method, which employs domain adversarial learning, effectively addresses the assignment bias and outperforms other methods for heterogeneous treatment effect estimation in the presence of interference.

Additional Authors: Jente Van Belle, Wouter Verbeke

Abstract number: 23

Date: Wednesday 9th April

Title: Method-of-Moments Inference for GLMs and Doubly Robust Functionals under Proportional Asymptotics

Presenter: Xingyu Chen

Affiliation: Shanghai Jiao Tong University

Topic: Machine Learning in Causal Inference

Abstract: In this paper, we consider the estimation of regression coefficients and signal-to-noise (SNR)ratio in high-dimensional Generalized Linear Models (GLMs), and explore their implications ininferring popular estimands such as average treatment effects in high dimensional observationalstudies. Under the "proportional asymptotic" regime and Gaussian covariates with known(population) covariance Σ, we derive Consistent and Asymptotically Normal (CAN) estimators ofour targets of inference through a Method-of-Moments type of estimators that bypasses estimationof high dimensional nuisance functions and hyperparameter tuning altogether. Additionally,under non-Gaussian covariates, we demonstrate universality of our results under certain additionalassumptions on the regression coefficients and Σ. We also demonstrate that knowing Σ is notessential to our proposed methodology when the sample covariance matrix estimator is invertible.Finally, we complement our theoretical results with numerical experiments and comparisons withexisting literature.

Additional Authors: Lin Liu, Rajarshi Mukherjee

Abstract number: 24

Date: Wednesday 9th April

Title: Prognostic scores and representation learning for causal effect estimation with weak overlap

Presenter: Oscar Clivio

Affiliation: University of Oxford

Topic: Machine Learning in Causal Inference

Abstract: Overlap, also known as positivity, is a key condition for modern causal machine learning. Many popular estimators suffer from high variance and become brittle when features strongly differ across treatment groups. This is especially challenging in high dimensions: the curse of dimensionality can make overlap implausible. Modern causalML methods typically address this issue only indirectly, leveraging dimension reduction or other representation learning that does not account for overlap. In the limit, such methods reduce features to a scalar, such as the prognostic score (i.e., the conditional counterfactual mean). Building on a venerable empirical literature, we argue that the prognostic score is an unreasonably effective dimension reduction approach, and is a promising default in otherwise complex settings. To show this, we first propose a class of feature representations called deconfounding scores, which preserve both identification and the target of estimation while also improving overlap; the propensity and prognostic scores are two special cases. We characterize the corresponding optimization problem in terms of controlling overlap under an unconfoundedness constraint. We then derive closed-form expressions for overlap-optimal representations under a broad family of generalized linear models with Gaussian covariates and show that this coincides with the prognostic score. We conduct extensive experiments to assess this behavior empirically.

Additional Authors: Alexander D'Amour, Alexander Franks, David Bruns-Smith, Avi Feller, Chris Holmes,

Abstract number: 25

Date: Wednesday 9th April

Title: Causal inference targeting a concentration index for studies of health inequalities

Presenter: Mohammad Ghasempour

Affiliation: Umeå University

Topic: Machine Learning in Causal Inference

Abstract: A concentration index, a standardized covariance between a health variable and relative income ranks, is often used to quantify income-related health inequalities. There is a lack of formal approach to study the effect of an exposure, e.g., education, on such measures of inequality. In this paper we contribute by filling this gap and developing the necessary theory and method. Thus, we define a counterfactual concentration index for different levels of an exposure. We give conditions for their identification, and then deduce their efficient influence function. This allows us to propose estimators, which are regular asymptotic linear under certain conditions. In particular, these estimators are $\sqrt{n}$-consistent and asymptotically normal, as well as locally efficient. The implementation of the estimators is based on the fit of several nuisance functions. The estimators proposed have rate robustness properties allowing for convergence rates slower than $\sqrt{n}$-rate for some of the nuisance function fits. The relevance of the asymptotic results for finite samples is studied with simulation experiments. We also present a case study of the effect of education on income-related health inequalities for a Swedish cohort.

Additional Authors: Xavier de Luna, Per E. Gustafsson

Abstract number: 26

Date: Wednesday 9th April

Title: Predicting Treatment Effects Across Contexts for Perturbational Biology

Presenter: Luka Kovačević

Affiliation: MRC Biostatistics Unit, University of Cambridge

Topic: Machine Learning in Causal Inference

Abstract: To effectively discover novel medicines, causal understanding of disease biology and cellular mechanisms is essential. By identifying these mechanisms, we can understand how diseases develop, manifest and how they can be treated in patients. Gene perturbation screens are often used to generate causal evidence for genetic drivers of disease; however perturbation screens are only viable in approximate cellular models, known as cell-lines, because CRISPR-based interventions induce significant stress in donor cells, making them unviable. While chemical interventions offer a less invasive alternative, they lack the specificity of CRISPR, making them less effective for elucidating causal pathways.Due to the scale of the human genome, and the contextual specificity of interventions, it is implausible to exhaust the space of possible experiments. Instead, perturbation prediction models have been proposed as an alternative to performing these experiments in the real world. Perturbation prediction models aim to predict the average treatment effects (ATEs) of genetic interventions and to enable in-silico simulations of interventions. However, existing methods often fail to generalise effectively to previously unseen cellular contexts.In this work, we introduce a novel paradigm for perturbation prediction that leverages chemical interventions as proxies for predicting CRISPR interventions in unseen cellular contexts. Using publicly available data, we demonstrate that, given a variety of chemical and CRISPR interventions across diverse cell lines, our approach can accurately predict CRISPR intervention effects in cell lines where only chemical interventions have been measured. Our findings show that the difficulty of predicting CRISPR ATEs varies based on the cell line lineage and the observed chemical intervention data. We evaluate a range of both causal and non-causal machine learning models previously proposed for perturbation prediction in this new problem setting. This work highlights the potential of chemical interventions as a bridge to improving CRISPR prediction models and expands the toolkit for studying causal mechanisms across complex cellular landscapes.

Additional Authors: Sach Mukherjee, John Whittaker, Thomas Gaudelet, Hagen Triendl

Abstract number: 27

Date: Wednesday 9th April

Title: Learning to rank treatment effects

Presenter: Myrto Limnios

Affiliation: Ecole Polytechnique Fédérale de Lausanne (EPFL)

Topic: Machine Learning in Causal Inference

Abstract: We explore the connection between learning to rank methods applied to treatment effects and optimal treatment allocation in the context of causal inference. An intrinsic characteristic of this setting is the unknown effect of a treatment to an individual, that usually results on strong model assumptions to solve the problem. In contrast to recent works on nonparametric methods for estimating the conditional treatment effect (CATE), we propose to optimize a summary statistic outputting the optimal rule, that statistically ranks with higher scores the individuals having positive treatment effects. This allows not only to consider that a specific treatment can be either beneficial or harmful depending on a threshold value, but to learn it for all individuals. Precisely, we propose to learn a scoring function maximizing empirical summaries of the Receiver Operating Characteristic (ROC) curve formulated as a linear rank statistic, and to use it for optimal treatment allocation. This framework is adapted to covariates values in multidimensional generic spaces, and we prove nonasymptotic generalization guarantees for the optimal empirical treatment scoring rule. This generic method provides great flexibility to the practitioner, by allowing various learning algorithms to solve the problem, namely bipartite ranking algorithms, as well as different summaries of the ROC curve adapted to the problem.

Abstract number: 28

Date: Wednesday 9th April

Title: Performance of Cross-Validated Targeted Maximum Likelihood Estimation

Presenter: Miguel-Angel Luque Fernandez

Affiliation: London School of Hygiene and Tropical Medicine, UK; University of Granada, Spain.

Topic: Machine Learning in Causal Inference

Abstract: Advanced methods for causal inference, such as targeted maximum likelihood estimation (TMLE), require fast convergence rates and the Donsker class condition for statistical inference. However, in situations where there is not differentiability due to data sparsity or near-positivity violations, the Donsker class condition is violated. In such situations, TMLE variance is inflated leading to poor coverage and conservative confidence intervals. Cross-validation of the TMLE algorithm (CVTMLE) is a simple, yet effective way to ensure efficiency, especially in settings where the Donsker class condition is violated.We investigated the performance of CVTMLE compared to TMLE in various settings. We contrasted several approaches to cross-validation of TMLE, including the original approach cross-validating the whole procedure proposed by Zheng and van der Laan (1), as well as the approach offered by Levy (2) suggesting to cross-validate the initial predictions of outcomes and treatment. We utilised the data-generating mechanism described in Leger et al. (3) to run a Monte Carlo experiment under different Donsker class violations. Then, we evaluated the respective statistical performances of TMLE and CVTMLE with different super learner libraries, with and without regression tree methods.We found that CVTMLE vastly improves confidence interval coverage without adversely affecting bias, particularly in settings with small sample sizes and near-positivity violations. Furthermore, incorporating regression trees using standard TMLE with ensemble super learner-based initial estimates increases bias and variance leading to invalid statistical inference.We show through simulations that CVTMLE is much less sensitive to the choice of the super learner library and thereby provides better estimation and inference in cases where the super learner library uses more flexible candidates and is prone to overfitting.1. Zheng Wenjing, Laan Mark van der. Asymptotic Theory for Cross-validated Targeted Maximum Likelihood Estimation. U.C. Berkeley Division of Biostatistics Working Paper Series. 20102. Levy Jonathan. An Easy Implementation of CV-TMLE. arXiv. 2018; https://doi.org/10.48550/arXiv.1811.045733. Léger M., Chatton A., Le Borgne F., Pirracchio R., Lasocki S., Foucher Y.. Causal inference in case of near-violation of positivity: comparison of methods. Biom J. 2022;64:1389-1403.

Additional Authors: Matthew Smith; Rachael Phillips; Camille Maringe

Abstract number: 29

Date: Wednesday 9th April

Title: Matching Methods for Difference-in-Differences with Multiple Time Periods: Evaluating the Equality of ATT Estimates Across Time

Presenter: Junho Jang

Affiliation: Department of Statistics, Seoul National University

Topic: Matching/weighting

Abstract: In observational studies with multiple time points, testing for homogeneous causal effects, such as the Average Treatment effect on the Treated (ATT), is crucial for evaluating treatment efficacy over time. This paper introduces a novel testing framework that accommodates arbitrary combinations of treatment initiation and post-treatment time points. For estimation, difference-in-differences (DID) is combined with matching to focus on post-treatment periods for treated units. Our testing framework involves two main steps. First, a confidence set for the common treatment effect is constructed to narrow the range of plausible parameters. Second, a randomization-based test is conducted within this confidence set to assess the equality of treatment effects. This approach extends multivariate location testing to partially matched sets. Furthermore, the relationship between matched set structure and test power is theoretically explored, providing insights to guide matching design in practice. To illustrate its application, we use this framework on Health and Retirement Study (HRS) data, testing and summarizing treatment equality across time periods.

Additional Authors: Yitae Kwon, Kwonsang Lee

Abstract number: 30

Date: Wednesday 9th April

Title: High-Dimensional Matching with Genetic Algorithms

Presenter: Hajoung Lee

Affiliation: Institute for Data Innovation in Science, Seoul National University, Republic of Korea

Topic: Matching/weighting

Abstract: Matching in observational studies is a widely used approach to estimate causal effects by obtaining treated and control groups with similar covariate distributions. Traditional methods for matching rely on distances between observations to form pairs. However, this process often faces challenges in high-dimensional and low-sample size settings due to the curse of dimensionality, where the concentration of distance makes it difficult to distinguish between observations. To address this problem, we propose a novel matching method using genetic algorithms, shifting the focus from individual-level to group-level distances. By optimizing the similarity of the high-dimensional joint distributions of covariates between treated and control groups, our method enables improved estimation of causal effects. This approach has several advantages: (1) it avoids dimension reduction or variable selection, thereby preserving the full scope of high-dimensional information without requiring additional modeling, and (2) it maintains transparency by not relying on outcomes, akin to traditional matching, and (3) it performs robustly in settings with low sample sizes, where traditional methods may struggle. Furthermore, our results show that the proposed method is competitive with existing approaches even in low-dimensional settings. Through extensive simulation studies and applications to real data, we validate the performance and provide practical guidance for the method, highlighting its potential as a powerful tool for causal inference in both high- and low-dimensional scenarios.

Additional Authors: Kwonsang Lee

Abstract number: 31

Date: Wednesday 9th April

Title: Evaluating time-specific treatment effects using randomization inference

Presenter: Sangjin Lee

Affiliation: Seoul National University

Topic: Matching/weighting

Abstract: This study develops a systematic approach for evaluating the effect of a treatment on a time-to-event outcome in a matched-pair study. While most methods for paired right-censored outcomes allow determining an overall treatment effect over the course of follow-up, they generally lack in providing detailed insights into how the effect changes over time. To address this gap, we propose novel tests for paired right-censored outcomes using randomization inference. We further extend our tests to matched observational studies by developing corresponding sensitivity analysis methods to take into account departures from randomization. Simulations demonstrate the robustness of our approach against various non-proportional hazards alternatives, including a crossing survival curves scenario. We demonstrate the application of our methods using a matched observational study from the Korean Longitudinal Study of Aging (KLoSA) data, focusing on the effect of social engagement on survival.

Additional Authors: Kwonsang Lee

Abstract number: 32

Date: Wednesday 9th April

Title: More Powerful Selective Inference in Propensity Score Analysis

Presenter: Sarah Pirenne

Affiliation: ORStat, KU Leuven

Topic: Matching/weighting

Abstract: We propose a method for post-selection inference in a causal inference setting with propensity scores and inverse probability weighted estimation. More specifically, we conduct inference for the effects of the covariates after variable selection by the least absolute shrinkage and selection operator (lasso). To obtain valid post-selection inference with control over the type I error rate, we use a selective inference approach, which conditions the inference on the event of model selection. By using a parametric programming method, we improve the statistical power of existing approaches to selective inference in this setting. Moreover, we allow data-based selection of the regularization parameter based on the inverse probability C_p (IPC_p) criterion. We illustrate the performance of our method with simulated data and an application on the Lalonde data (Lalonde, 1986).

Additional Authors: Gerda Claeskens and Yoshiyuki Ninomiya.

Abstract number: 33

Date: Wednesday 9th April

Title: Selection bias due to omitting interactions from inverse probability weighting

Presenter: Liping Wen

Affiliation: University of Bristol, United Kingdom

Topic: Matching/weighting

Abstract: The estimated causal effect of an exposure on an outcome might be biased if the datasets are subject to selection due to non-random participation, dropout, or intermittent measures. Inverse probability weighting (IPW) is often used to adjust for selection bias, usually using a simple logit/probit model without interactions. However, our recent work shows that for a given analysis, the size of the selection bias depends on the interaction between the exposure and outcome in their effect on selection. This implies that it may be important to include interaction terms in the IPW model. Via simulations and a real-data application we compare the performance of IPW with and without interaction terms. The simulation study shows that IPW including interaction terms gives less biased estimates than IPW without interactions in all scenarios studied. Importantly, IPW using a logistic model with no interaction terms often gives estimates very close to the complete case analysis (CCA) – perhaps giving false reassurance that results are robust to selection bias. In a real-data application, we use data from the Understanding Society study to investigate the effect of unemployment on sleep duration. IPW including interaction terms suggests that unemployment reduces sleep duration by around 24 minutes (9, 38), compared to 27 minutes (14, 40) for IPW without interactions, and 31 (19, 44) minutes for CCA. We recommend that applied researchers using IPW for selection bias routinely incorporate interactions in their weighting model.

Additional Authors: Apostolos Gkatzionis, Kate Tilling, Rosie Cornish, Rachael Hughes

Abstract number: 34

Date: Wednesday 9th April

Title: Estimating Average Treatment Effect via Marginal Outcome Density Ratio

Presenter: Linying Yang

Affiliation: University of Oxford

Topic: Matching/weighting

Abstract: Doubly robust estimators, such as AIPW, offer the advantage of providing 'two chances' to perform estimation correctly and still obtain a consistent estimator. However, due to inverse probability weighting by the propensity score, these estimators can suffer from practical positivity violation, where some covariates predict the treatment so well that our weights become extremely large; this inflates the efficiency bound and estimation variance. This leads to the concept of the marginal density ratio. Instead of manually evaluating propensity scores or selecting features in the pre-treatment covariate space, we shift our focus to the outcome space, allowing the observed outcomes to determine which information should be included. In this paper, we introduce the Marginal outcome density Ratio estimator (MR) and the Augmented Marginal outcome density Ratio estimator (AMR); we demonstrate the advantages of these estimators in filtering necessary information to obtain treatment effects. We argue in this paper that, using this information filtering, MR and AMR are able to estimate the average treatment effect more effectively in small samples than their direct counterparts, IPW and AIPW. We also give examples on its contribution to sparsity condition of ATE estimation in the high-dimensional context.

Additional Authors: Robin Evans

Abstract number: 35

Date: Wednesday 9th April

Title: Causal discovery for multi-cohort studies

Presenter: Christine Bang

Affiliation: University of Copenhagen

Topic: Network discovery

Abstract: The availability of multiple overlapping cohort datasets enables us to learn causal pathways over an entire lifespan. Evidence of such pathways may be highly valuable, e.g. in life course epidemiology. No previous causal discovery methods tailored to this framework exist. We show how to adapt an existing causal discovery algorithm for overlapping datasets to account for the time structure embedded in cohort data. In particular, we show that this strengthens the method in multiple aspects.Constraint-based causal discovery methods recover causal structures from (conditional) independencies. Multiple causal structures may induce the same dependence structure, and form an equivalence class. Without additional, stronger assumptions, it is usually not possible to recover more than the equivalence class; i.e. we cannot identify all causal directions. Moreover, when combining multiple datasets, if some variables are never measured jointly their (conditional in-)dependence is by construction unknown. Then, we cannot even identify the equivalence class. Hence, constraint-based causal discovery for multiple datasets suffers from two types of obstacles for identification. Time structured data induces a partial causal ordering of the variables, which we refer to as tiered background knowledge. It is easy to see that tiered background knowledge improves the (partial) identifiability of causal directions. Additionally, we show that tiered background knowledge also improves the (partial) identifiability of the equivalence class, which is not trivial. We provide theoretical results on the informativeness as well as theoretical guarantees of the algorithm. Finally, we provide detailed examples that illustrate how and when tiered background knowledge increases the level of informativeness.

Additional Authors: Vanessa Didelez

Abstract number: 36

Date: Wednesday 9th April

Title: Nonparametric conditional independence testing with mixed data

Presenter: Luca Bergen

Affiliation: Leibniz Institute for Prevention Research and Epidemiology

Topic: Network discovery

Abstract: Causal discovery is becoming increasingly important in epidemiology as well as other scientific fields, e.g., to guide the selection of adjustment sets for causal effect estimation or the identification of intervention targets for, say, preventive measures. Of special importance are constraint-based methods for causal discovery, as they are relatively flexible and can allow for latent variables. However, the vast majority of epidemiological data have mixed measurement scales, e.g., continuous, nominal and ordinal variables. Yet, crucially, there exists no conditional independence (CI) test for mixed data that can be considered a good, or generally accepted, default. The main parametric test is based on the Conditional Gaussian distribution (Lauritzen & Wermuth, 1989), which is often not justified by prior distributional knowledge, while nonparametric tests have not been comprehensively compared and validated.Our aim is to give some guidance on the choice of CI test under a wide range of data-generating processes including interactions, nonlinear basis functions and post-nonlinear distortion. We therefore investigate and compare the utility of a wide variety of recently proposed (nearly) nonparametric CI tests for mixed data in causal discovery. This includes methods based on residual testing of supervised ML models (Shah & Peters, 2020; Ankan & Textor, 2023), feature importance scores as used in explainable ML (Dai et al., 2024; Tansey et al., 2022; Watson & Wright, 2021), as well as kernel methods (Zhang et al., 2011; Strobl et al., 2019; Bach & Jordan, 2002). We contrast the theoretical principles underlying these tests; we then discuss their strengths and weaknesses, statistical properties , and compare their (finite sample) performance in causal discovery, as quantified by precision and recall of adjacencies and v-structures as well as runtime, against a parametric default test in a simulation study. While our work will not be able to provide a universally valid guidance regarding the choice of CI test, it is an important step forward to help users of constraint-based causal discovery methods make a better-informed decision and motivate further comparisons.

Additional Authors: Vanessa Didelez

Abstract number: 37

Date: Wednesday 9th April

Title: Scalable Constraint-Based Causal Discovery with Multiple Imputation for Incomplete Data

Presenter: Frederik Fabricius-Bjerre

Affiliation: University of Copenhagen

Topic: Network discovery

Abstract: Constraint-based causal discovery methods, such as the PC-algorithm, can be used to recover causal structures from data by leveraging conditional independence tests. But what do you do in the presence of missing data? In regression analyses, multiple imputation offers an effective strategy for handling missing entries. However, previous research has found that for causal discovery, multiple imputations do not always perform better than simplistic test-wise deletion strategies even in cases where theoretical results suggest it should. Fully exploring the practical potential of combining causal discovery with multiple imputation has however been infeasible due to the relatively large computation time of both procedures. It is thus not clear how to best handle missing information when using causal discovery in practice.To address this knowledge gap, we present research integrating multiple imputation with both the PC-algorithm and the temporal PC-algorithm, examining the balance between computational efficiency and the quality of the estimated causal structures. We derive theoretical upper computational bounds for the PC and TPC algorithms with multiple imputation. Furthermore, we provide a GPU implementation of the (temporal) PC-algorithm with multiple imputations for Gaussian data. This reduces runtime by a factor of 92.8, and thus makes it possible to conduct much more comprehensive investigations into the combination of multiple imputations and PC.In such investigations, we explore the utility of varying both the number of imputations and the degrees of freedom approximations for pooled conditional independence tests in multiple imputation for linear Gaussian structural causal models with diverse network structures. Crucially, results indicate that increasing the number of imputations up to even 1000 imputations consistently enhances the quality of causal structure estimation for both the PC-algorithm and temporal PC-algorithm on incomplete data. Hence, we conclude that although computationally expensive, a large number of imputations may be worthwhile for causal discovery. And by using the new GPU implementation, it is now also much more feasible to do so in practice.

Abstract number: 38

Date: Wednesday 9th April

Title: When Causal Discovery Errs: Inconsistency Detection for PC-like Algorithms

Presenter: Sofia Faltenbacher

Affiliation: TU Dresden

Topic: Network discovery

Abstract: The PC algorithm is the foundation of a range of constraint-based causal discovery algorithms such as PC-stable, conservative-PC, FCI, and PCMCI, to name a few. Under certain assumptions, these algorithms learn correct equivalence classes of the ground truth causal graph asymptotically. However, real-world data will often violate the conditions necessary for the soundness of a PC-like algorithm, beginning with the unavoidable limitation of finite samples. Therefore, it is important to analyze and quantify how assumption violations and sampling errors can change the results of the algorithms. In this talk, I classify error types and their manifestations in the output graph with easy-to-understand illustrative examples. I discuss the importance of the subset of errors that leave their imprint as inconsistencies between the conditional independence tests performed by the PC-like method and the separation statements implied by the output graph. Finally, I propose several scores to quantify the internal inconsistencies of PC-like algorithms.

Additional Authors: Jonas Wahl, Rebecca Herman, Jakob Runge

Abstract number: 39

Date: Wednesday 9th April

Title: Small Open-Source LLMs for Causal Discovery [chatGPT-generated by local organizers as no title was provided]

Presenter: Louis Hernandez

Affiliation: Craft AI / INSA Rouen

Topic: Network discovery

Abstract: This study investigates the potential of small open-source large language models (LLMs) for causal discovery tasks, a domain currently dominated by proprietary models like GPT-4. We introduce a constrained generation method to enforce structured, interpretable outputs, effectively eliminating formatting errors, and evaluate the impact of few-shot learning to determine if in-context examples can improve causal inference accuracy. Our experiments on the ASIA and CHILD datasets reveal that while constrained generation enhances output consistency, it does not significantly improve F-scores. Few-shot learning provides only modest gains, underscoring the limitations of smaller LLMs in modeling complex causal relationships. These findings highlight the unique challenges smaller models face in causal discovery and suggest that future hybrid approaches—integrating open-source LLMs with traditional data-driven causal inference algorithms—may offer a more effective path to accurate and reliable causal discovery in real-world applications.

Additional Authors: Alessandro Leite, Matthieu Boussard

Abstract number: 40

Date: Wednesday 9th April

Title: Score-Based Causal Discovery with Temporal Background Knowledge

Presenter: Tobias Ellegaard Larsen

Affiliation: Section of Biostatistics, University of Copenhagen

Topic: Network discovery

Abstract: Causal conclusions from statistical studies often rely on causal assumptions represented by DAGs. In fields like epidemiology these assumptions are most widely based on prior knowledge, making the process confirmatory and limited in its ability to reveal novel causal pathways. Causal discovery offers a data-driven alternative to the traditional approach for uncovering causal relationships.We present our implementation of an extension to the Greedy Equivalence Search (GES) algorithm, which takes in temporal background knowledge in an efficient way. We show results stating that the desirable theoretical properties of GES, such as correct equivalence class estimation in the large-sample limit, are inherited by the extension. While arguably more important, the algorithm also performs well for finite samples. We will observe how the algorithm, given data such as cohort data, results in estimates with high recall and precision for both sparse and dense underlying mechanisms, making it a useful tool in real-world applications.Our focus will be centered around score-based causal discovery with temporal background knowledge, but we will relate it to a constraint-based alternative. This comparison will highlight the trade-offs between using the one approach or the other.

Additional Authors: Anne Helby Petersen, Clausk Thorn Ekstrøm

Abstract number: 41

Date: Wednesday 9th April

Title: Full pipeline for causal inference on limited sample size

Presenter: Tobias Strømgren

Affiliation: University of Copenhagen, Section of Biostatistics

Topic: Network discovery

Abstract: Causality research is broadly divided into two key areas: causal discovery, which focuses on identifying the causal data-generating mechanism, and causal inference, which quantifies the causal effects between variables and their associated uncertainties. We present methods bridging these two areas and extend existing methods to create a more comprehensive pipeline that enables causal interpretations through a fully data-driven approach, starting from raw data.Many causal inference methods rely on the assumption that the data-generating process aligns with a known Directed Acyclic Graph (DAG). These DAGs are typically provided by domain experts; however, variations in expert opinions—and the resulting discrepancies in their proposed DAGs—can lead to inconsistent and non-robust inferences.Recent advancements show that causal effects can be bounded without fully specifying the data-generating process, by relaxing assumptions to the Markov equivalence class or restricted Markov equivalence class of the true DAG (Perković et al., 2018; Fang & He, 2020). However, with finite data and without strong parametric assumptions, causal discovery algorithms may encounter conflicts arising from statistical errors. These conflicts can result in graphs that fall outside the scope of any Markov equivalence class, such as graphs containing cycles. Current causal inference methods struggle to accommodate such graphs.We will present advantages and limitations of utilizing causal discovery combined with causal inference. This includes an extension of the Intervening with DAG Absent (IDA) algorithm introduced by Maathuis et al. (2009), designed to handle graphs beyond the restricted Markov equivalence class. This method bridges the gap between causal discovery and inference, making a fully data-driven approach more achievable. Our findings underscore the challenges of relaxing assumptions about the data-generating process, but they also demonstrate how and when causal reasoning remains feasible without stringent and often unrealistic/unknown structural assumptions.Results from simulation studies - including both large sample limits and finite sample scenarios - will be presented alongside applications to real-world data.

Additional Authors: Anne Helby Petersen, Claus Thorn Ekstrøm

Abstract number: 42

Date: Wednesday 9th April

Title: Measuring similarity of causal graphs

Presenter: Jonas Wahl

Affiliation: German Research Centre for Artificial Intelligence (DFKI)

Topic: Network discovery

Abstract: To evaluate the quality of methods that infer causal graphs from data, assumptions and expert knowledge, it is necessary to measure how similar two causal graphs are to each other. At the same time, there are multiple ways to operationalize similarity of causal graphs by means of their predictive, structural and interventional implications. In this talk, we will present several recently developed metrics for causal graph comparison, discuss their strengths and limitations as well as the settings in which they can or can't be applied.

Additional Authors: Jakob Runge

Abstract number: 43

Date: Wednesday 9th April

Title: Selection bias of cause-specific hazard ratios: the impact of competing events

Presenter: Mari Brathovde

Affiliation: Oslo University Hospital

Topic: Other

Abstract: Competing risks generalize standard survival analysis of a single, often composite outcome when interest lies in the different causes of the event. In the presence of heterogeneity, the complex causal interpretation of the hazard ratio for all-cause mortality is well-known and has been formalized. Yet the current recommendation in epidemiology is to use cause-specific hazard ratios when aiming to understand etiology in a competing risk setting. In this work, we formalize how observed cause-specific hazard ratios evolve and deviate from the (conditional) causal effect of interest in the presence of heterogeneity of the hazard rate of unexposed individuals (frailty) and heterogeneity in effect (individual modification). We show that the presence of a competing event can amplify the selection bias of the all-cause mortality setting, as it introduces selection on the frailties and effect modifiers associated with both the event of interest and the competing event.Furthermore, we show that the size and sign of the bias in the cause-specific hazard ratios depend on the prevalence of, and the treatment's effect on, the competing event. Consequently, the cause-specific hazard ratio suffers from both selection bias inherent to all-causemortality hazard ratios and dependence on the competing event, complicating its causal interpretation and limiting its suitability for addressing etiological questions without relying on untestable assumptions.We provide illustrative examples using frailties from the family of power variance function (PVF) distributions, along with categorical effect modifiers (harmful, beneficial, or neutral). The PVF family of frailties yields convenient analytical expressions for the observed cause-specific hazard ratios, enabling a clear separation of the selection bias introduced by different events. This approach allows for a straightforward evaluation of the impact of treatment effects and event prevalence on the bias using simple monotonicity principles. The numerical examples include settings with crossover of the cause-specific hazard rates between exposed and non-exposed individuals, which would not occur without competing events. This work highlights the importance of employing more appropriate estimands in a competing risk setting, such as marginal cumulative incidences.

Additional Authors: Morten Valberg, Hein Putter, Richard Post

Abstract number: 44

Date: Wednesday 9th April

Title: Decision Analytical Models as Causal Models

Presenter: Maurice Korf

Affiliation: Erasmus University Medical Center (EMC)

Topic: Other

Abstract: Decision Analytical Models as Causal ModelsDecision analytical models are widely used to compare the costs and benefits of health interventions through synthesizing evidence. These models simulate individuals or cohorts under each intervention to estimate and compare the cost-effectiveness. As a result, the most cost-effective health intervention can be identified and used to inform policy.These cost-effectiveness analyses are, however, only informative if they accurately approximate the causal impact of each decision. Although cost-effectiveness models are inherently causal, the literature lacks explicit causal language, and causal assumptions are often not evaluated or even mentioned. In this work, we illustrate that the incremental cost-effectiveness ratio (ICER), the typical estimand targeted in cost-effectiveness analyses, is a causal quantity. Specifically, we show that the ICER is a function of multiple causal estimands, as a single intervention is typically modelled to influence multiple distinct, sequential outcomes. This, however, introduces causal considerations that are not always straightforward. Using a simple decision tree model with a single decision node and two descendant nodes, we illustrate that even in this basic structure, unconventional causal estimands—specifically, counterfactuals conditioned on counterfactuals—are required for the input parameters.We establish the correct estimands for the input parameters in these models and systematically outline the necessary assumptions to unbiasedly estimate and identify the target estimand, distinguishing between assumptions related to the decision model structure and those necessary for identifying the input parameters. We particularly emphasize the importance of external validity, as the decision model can be viewed as a data-fusion problem.In summary, we present decision analytical models as causal models, highlighting that the more a decision model deviates from its underlying causal structure —whether in design or the parameters used in the model —the more likely it is to result in a sub-optimal decision with all the consequences this entails.

Additional Authors: Myriam Hunnik, Richard Post, Jeremy Labrecque

Abstract number: 45

Date: Wednesday 9th April

Title: Discussion on treatment effect waning

Presenter: Eni Musta

Affiliation: University of Amsterdam

Topic: Other

Abstract: Understanding how the causal effect of a treatment evolves over time, including the potential for waning, is important for informed decisions on treatment continuation or repetition. For example, waning vaccine protection influences booster dose recommendations, while cost-effectiveness analyses require accounting for long-term efficacy of treatments. However, there is no consensus on the methodology to assess and account for treatment effect waning. Naïve comparisons of hazard functions, even in randomized controlled trials, can lead to misleading causal conclusions due to inherent selection bias. Although comparing survival curves is recommended as a safer measure of causal effect, it only represents a cumulative effect over time and does not address treatment effect waning. We explore recent formulations of causal hazard ratios, based on the principal stratification approach or the controlled direct effect. These causal hazard ratios cannot be identified without strong untestable modeling assumptions, but bounds can be derived accounting for unobserved covariates and one could try to use them to identify treatment effect waning. However we discuss that the concept of waning is itself not uniquely defined. Through examples, we illustrate that an increase in causal hazard ratios towards one does not necessarily mean the protective effect of the treatment is fading. Furthermore, the same survival functions may correspond to scenarios with or without waning, highlighting the need for a better understanding and guidance of what can and cannot be inferred from the observed data.

Additional Authors: Joris Mooij

Abstract number: 46

Date: Wednesday 9th April

Title: A Diagnostic to Find and Help Combat Positivity Issues— with a Focus on Continuous Treatments and Modified Treatment Policies

Presenter: Katharina Ring

Affiliation: University of Munich (LMU)

Topic: Other

Abstract: The positivity assumption is a central assumption for the identification of a causal effect, yet is rarely discussed, especially in conjunction with continuous treatments. One general recommendation for dealing with a violation is to change the estimand. However, an applied researcher is faced with two problems:First, how can she tell whether there is an actual positivity violation given her estimand of interest, preferably without having to estimate a model first? This is especially challenging for continuous treatments or longitudinal data, where simply checking cross tables is not possible.Second, if she finds a problem with positivity, how should she change her estimand in order to arrive at a meaningful estimand which does not face the same issues with positivity? Ideally, this new estimand should also be as close as possible to her original research question.We suggest a novel diagnostic which allows the researcher to answer both questions by providing insights into how well a prediction for a certain estimand can be made using the data at hand. The diagnostic breaks down the multi-dimensional information about whether observed data points are near the locations where we aim to make predictions into one dimension by using a proximity measure. In other words it quantifies the number of nearby data points, weighted by their proximity to the prediction point. This results in a count of "effective data points" for each prediction problem, which allows for identifying observations with little support, for a given intervention strategy. This information can be used to adjust the estimand by, for example, changing the intervention for these individuals.Modified Treatment Policies are a large class of estimands from which one might choose an alternative to combat positivity issues. We provide a simulation study on the general behaviour of such estimands at different levels of positivity violations and show how the diagnostic helps understand where bias is to be expected. We illustrate our proposed diagnostic and resulting choice of estimand in a pharmacoepidemiological study, based on data from CHAPAS-3, a trial comparing different treatment regimens for children living with HIV.

Additional Authors: Michael Schomaker

Abstract number: 47

Date: Wednesday 9th April

Title: Post-treatment problems: What can we say about the effect of a treatment among sub-groups who (would) respond in some way?

Presenter: Tanvi Shinkre

Affiliation: University of California Los Angeles

Topic: Other

Abstract: Investigators are often interested in how a treatment affects an outcome for units responding to treatment in a certain way. We may wish to know the effect among units that, for example, meaningfully implemented the intervention, passed an attention check, or survived to an endpoint of interest. Simply conditioning on the observed value of the relevant post-treatment variable introduces problematic biases. Further, assumptions such as "no unobserved confounding" (of the post-treatment mediator and the outcome) or of "no direct effect" (of treatment on outcome) required of several existing strategies are typically indefensible. We propose the Treatment Reactive Average Causal Effect (TRACE), which we define as the total effect of the treatment in the group that, if treated, would realize a particular value of the relevant post-treatment variable. Given the total effect of treatment, and by reasoning about the treatment effect among the "non-reactive" group, we can identify and estimate the range of plausible values for the TRACE. We discuss this approach and its connection to existing estimands and identification strategies, then demonstrate its use with two applications: (i) a community-policing intervention in Liberia, for which we estimate the effect of intervention among locations where the project was meaningfully implemented, and (ii) a field experiment studying how in-person canvassing affects support for transgender rights, for which we estimate the effect of canvassing among participants whose feelings towards transgender people become more positive.

Additional Authors: Chad Hazlett, Nina McMurry

Abstract number: 48

Date: Wednesday 9th April

Title: Quantification of Vaccine Waning as a Challenge Effect

Presenter: Matias Janvin

Affiliation: University of Oslo

Topic: Sensitivity Analysis/bounds

Abstract: Knowing whether vaccine protection wanes over time is important for health policy and drug development. However, quantifying waning effects is difficult. A simple contrast of vaccine efficacy at two different times compares different populations of individuals: those who were uninfected at the first time versus those who remain uninfected until the second time. Thus, the contrast of vaccine efficacy at early and late times can not be interpreted as a causal effect. We propose to quantify vaccine waning using the challenge effect, which is a contrast of outcomes under controlled exposures to the infectious agent following vaccination. We identify sharp bounds on the challenge effect under nonparametric assumptions that are broadly applicable in vaccine trials using routinely collected data. We demonstrate that the challenge effect can differ substantially from the conventional vaccine efficacy due to depletion of susceptible individuals from the risk set over time. Finally, we apply the methods to derive bounds on the waning of the BNT162b2 COVID-19 vaccine using data from a placebo-controlled randomized trial. Our estimates of the challenge effect suggest waning protection after 2 months beyond administration of the second vaccine dose.

Additional Authors: Mats J. Stensrud

Abstract number: 49

Date: Wednesday 9th April

Title: Expressing Cost of Causal Assumptions Through Partial Identification

Presenter: Jakob Zeitler

Affiliation: University of Oxford

Topic: Sensitivity Analysis/bounds

Abstract: This work is a first step towards a cost-sensitive perspective on causal modelling, evaluating the cost, expected results and risks associated with causal assumptions, through the lens of partial identification. It highlights the need for more careful specification of causal assumptions, and evaluation of causal inferences, in the context of real-world financial and human costs.Causal Inference is a language to express assumptions about the world that yield insights for decision making. Which HIV-drug should I prescribe to a patient for symptom treatment? Which discount code should I send to a customer to prevent churn?The majority of causal assumptions are empirically hard to verify, with results from purely observational studies being harder to justify compared to partial or complete experimental studies, such as RCTs. Partial Identification calculates lower and upper bounds on the true causal effect, with stronger assumptions yielding tighter bounds, sometime collapsing to full point-identification as experimental data would deliver.These stronger assumptions come at direct and indirect costs, mainly financial and human. Statistics as a discipline, in general, tries to inform decision makers by providing a range of plausible models that vary in strength and sometimes yield contradicting recommendations. It is often then up to the subject matter expert to decide which model is most plausible given the data, often trading-off the benefit on subgroup for treatment with the negative impact on another subgroup (Look AHEAD trial, Baum et al, 2017).On one side, experimental studies provide strong and reliable evidence, but are costly to obtain, e.g. expensive clinical trials are required for drug efficacy and preventing toxic and potentially lethal treatments, i.e. avoiding human costs. On the other side, observational studies become actionable only once ever stronger and hard to defend assumptions are made. With partial identification, the strength of these assumptions can be explicated as a range, often smoothly trading-off strength with actionable causal effect.This work discusses the central trade-off between how informative a causal inference is and the strength of the assumptions required to attain those inferences. It will take on the lens of partial identification and try to answer: What are the costs of causal assumptions?

Abstract number: 50

Date: Wednesday 9th April

Title: Target trial emulation in biobank data: estimating the effect of cholesterol-lowering therapy and PRS in the Estonian Biobank

Presenter: Saskia Kuusk

Affiliation: University of Tartu

Topic: Time-varying confounding/target trials

Abstract: For estimating the effects of a treatment on an outcome, Randomized Control Trials (RCTs) are acknowledged to be the gold standard. Population-based biobanks, however, offer a different, more cost- and time-effective approach to use observational data from Electronic Health Records. Although such data is not collected for research purposes and lacks random treatment assignment, it is still possible to obtain the same estimates as one would get from an RCT. We will demonstrate the first steps of such Target Trial Emulation (TTE) method on Estonian Biobank data to estimate the effect of cholesterol-lowering therapy and a Polygenic Risk Score (PRS) on cardiovascular disease. This includes specifying the trial components as they would appear in a hypothesised RCT and how to emulate them in our observational data. By conducting a TTE analysis in a biobank context, we aim to identify the protective effect of cholesterol-lowering therapy on cardiovascular diseases, validate the effect of PRS, and identify possible interactions between PRS and treatment. As we advanced in our analysis, we addressed several common pitfalls inherent in TTE analyses, such as classical immortal-time bias and the omission of key risk factors. Addressing these issues is crucial for accurately estimating the true causal effect of interest. To tackle the issue of immortal-time bias, we have implemented additional measures and are employing sequential TTE to mitigate this bias effectively.

Additional Authors: Lili Milani, Anastassia Kolde, Krista Fischer

Abstract number: 51

Date: Wednesday 9th April

Title: A Graphical Approach to State Variable Selection in Off-policy Learning

Presenter: Joakim Blach Andersen

Affiliation: University of Cambridge

Topic: Time-varying confounding/target trials

Abstract: Sequential decision processes are widely studied across many areas of science. A key challenge when learning policies from observational data - a practice generally referred to as off-policy learning - is how to "identify' ' the impact of a policy of interest when the observed data is not randomized. This issue has mainly been adressed in two prominent research areas: dynamic treatment regimes (DTRs) and offline reinforcement learning (RL). While the DTR literature primarily focuses on controlling for confounding in settings with short decision horizons, the RL literature focuses on dimension reduction in controlled environments such as games. This gap between the two literatures has limited the wider application of off-policy learning to many real-world problems.Using acyclical directed mixed graph (ADMGs), we provide a set of graphical criteria in a general decision process that encompasses both DTRs and MDPs, that guarantee identification. We discuss the implicit and explicit causal assumptions made in the DTR and RL literature, and how our results relate to them. Finally, we present a simulation study for the dynamic pricing problem encountered in container logistics, and demonstrate how violations of our assumptions can lead to suboptimal policies.

Additional Authors: Qingyuan Zhao

Abstract number: 52

Date: Wednesday 9th April

Title: Theory and Application of longitudinal causal inference for comparative effictiveness in diabetes research

Presenter: Thomas Gerds

Affiliation: University of Copenhagen

Topic: Time-varying confounding/target trials

Abstract: Emulated trials, combined with longitudinal causal inference, have the potential to provide valuable new insights into real-world drug efficacy. To establish credibility among practitioners, the initial milestone of our ongoing pharmacoepidemiology research on diabetes mellitus involves comparing results from randomized controlled trials (RCTs) with findings derived from European registry data. For both analyses, we employ a framework based on longitudinal targeted minimum loss-based estimation (LTMLE).The next phase focuses on evaluating the comparative effectiveness of drugs in populations and subgroups that would not have met the inclusion criteria of the original trials. For the initial step, we identify individuals in the registries who meet the eligibility criteria of major RCTs and re-analyze data from these trials. However, our efforts to apply theoretical knowledge from causal inference and statistical learning have revealed numerous pitfalls and challenges, highlighting the need for further methodological and software development for LTMLE.Key topics addressed in this work include the definition of estimands, the discretization of continuous-time treatment and covariate data, and the application of superlearning. Specific challenges include handling treatment holidays and rescue medications in RCTs, as well as accounting for polypharmacy in registry data. These findings underscore the complexity of translating theoretical frameworks into practical tools for pharmacoepidemiology research.

Additional Authors: Kathrine Kold Sørensen

Abstract number: 53

Date: Wednesday 9th April

Title: Model-free estimands for target trial analysis

Presenter: Edoardo Efrem Gervasoni

Affiliation: Ghent University

Topic: Time-varying confounding/target trials

Abstract: The target trial framework is a powerful methodology to estimate causal effects in observational studies by emulating randomized controlled trials. It addresses common biases in observational data, such as confounding and selection bias, by conceptualizing a sequence of hypothetical trials initiated at different time points. To increase precision, information is pooled across trials, typically under the assumption that the treatment effect is constant over time and across individuals. However, this can create ambiguity about the causal question when the effects vary or, as frequently happens, the population observed in some emulated trials systematically differs from the target population. Further challenges come when effects are parametrized using non-linear regression models (e.g. logistic regression) where non-collapsibility can create issues of misspecification when pooling different populations.To address these challenges, this project introduces an assumption-lean strategy for target trial analysis, focussing on the choice of the estimand, rather than the choice of a model. This ensures that the analysis' aim is unequivocal regardless of model misspecification, and that uncertainty assessments reflect only information available in the data. Our proposal consists of several estimands, each related to different data structures and addressing different aspects of the patient population that may be of interest to researchers or decision-makers. For these estimands, corresponding estimators have been developed by the use of G-computation and inverse probability weighting. Applications on simulations and real data on antimicrobial de-escalation in an intensive care unit setting demonstrate the advantages of the proposed methodology over traditional techniques, offering greater clarity and reliability in causal effect estimation.

Additional Authors: Oliver Dukes, Stijn Vansteelandt

Abstract number: 54

Date: Wednesday 9th April

Title: Structural Nested Models in Target Trial Emulation

Presenter: Fuyu Guo

Affiliation: Harvard University

Topic: Time-varying confounding/target trials

Abstract: Target trial emulation is a popular method for estimating effects of treatment regimes from observational data. In the emulation, new trials, indexed by time, are initiated at fixed intervals. A subject participates in every trial for which eligibility criteria are met. Current methods treat each time-specific trial separately. For instance, for a trial comparing the regimes "always" versus "never" treat from initiation at t onwards, it is common to fit a hazard or risk ratio (RR) model that includes a treatment indicator and its potential confounders. Subjects are censored if they later change treatment, with inverse probability weighting to adjust for the censoring. If most subjects change treatment, the estimates will be inefficient. In this paper we propose more efficient estimators by introducing regime-specific structural nested target trial emulation models (SNTTEM). Given a regime, a SNTTEM imposes parametric models for all time-specific blip functions of the eligible subjects and leaves those for the ineligible unrestricted. A time-specific blip function quantifies on a mean scale the effect of initiating the regime at a time t versus one period later, as a function of past history. The intersection of all the earlier time-specific RR models constitutes a single SNTTEM with regime "always take the treatment that one took last time". We show that SNTTEM can be fitted using g-estimation, a method that censors less and is more efficient than current methods.

Additional Authors: Oliver Dukes; Mats Julius Stensrud; James Robins

Abstract number: 55

Date: Wednesday 9th April

Title: Improving inference procedures in the R package TrialEmulation with resampling methods

Presenter: Juliette Limozin

Affiliation: MRC Biostatistics Unit, University of Cambridge

Topic: Time-varying confounding/target trials

Abstract: Randomised controlled trials are the gold standard for estimating causal treatment effects, but their practical limitations often necessitate alternative approaches using observational data, e.g. from electronic health records. The target trial emulation framework (Hernán and Robins, 2016) provides a formal methodology for estimating causal effects from such data. The R package TrialEmulation, developed by Gravestock et al. (2024), supports this framework by enabling users to emulate a single or a sequence of target trials with time-to-event outcome. The package offers data preparation, inverse probability weighting to handle treatment switching and dependent censoring, marginal structural model fitting, and marginal risk difference estimation. Currently, confidence intervals (CIs) in TrialEmulation are constructed solely using the sandwich variance estimator, which may have limitations in small-sample or complex-data settings.In this poster, we present enhancements to TrialEmulation motivated by Limozin et al. (2024), who compared methods for constructing CIs in sequential trial emulation. They considered nonparametric bootstrap, linearised estimating function (LEF) bootstrap, and jackknife resampling as alternatives to the sandwich variance estimator-based method. Their findings highlighted the benefits of LEF bootstrap, which provided better coverage in data scenarios with small/moderate sample sizes, low/medium event rates and low/medium treatment prevalence. By integrating these resampling methods into TrialEmulation, we provide researchers with various tools for constructing CI in different settings.The poster will detail the theoretical foundations of these resampling methods, their implementation in the package, and their application to simulated data. These enhancements improve the inference procedures in TrialEmulation and broaden the package's utility for researchers employing target trial emulation.

Additional Authors: Isaac Gravestock, Shaun Seaman, Li Su

Abstract number: 56

Date: Wednesday 9th April

Title: Analysis of Longitudinal Causal Effects Using Joint Logit Models

Presenter: Ji Luo

Affiliation: Zhejiang University of Finance and Economics

Topic: Time-varying confounding/target trials

Abstract: Background: Estimating unbiased causal effects of time-varying treatments in observational longitudinal data is challenging, particularly when influenced by time-dependent and unobservable confounders. Most methods assume away unobservable confounders, an unrealistic assumption that complicates causal estimation, especially with non-randomly missing data.Methods: This paper introduces a method combining joint Logit models (JLM) with the G-computation formula to address causal inference for time-varying treatments. It uses random effects as proxies for unobservable confounders and extends the JLM to handle monotone missing data and death truncation. The method is validated through simulation studies and applied to a real-data analysis of 86 schizophrenia patients' recovery post-hospitalization in Madras.Results: The simulation study shows the proposed JLM's superiority in handling binary outcomes and unobservable confounding factors. The real-data analysis reveals that both gender and age have negative causal effects on recovery speed, with age showing a more significant impact. The JLM outperforms Longitudinal Random Effects Logit Model (GLM) and Fixed Effects Logit Model (FE) models in estimating causal effects, especially with missing data and time-dependent confounding.Conclusions: The JLM effectively addresses unobservable confounding and missing data in causal inference of time-varying treatments. The application to schizophrenia recovery data highlights the significant negative impact of age on recovery speed, demonstrating the JLM's superiority in providing accurate causal inferences over traditional models.

Additional Authors: Ji Luo

Abstract number: 57

Date: Wednesday 9th April

Title: A visual guide to g-estimation to handle symptomatic treatment at multiple timepoints in a randomized controlled trial

Presenter: Baldur Magnusson

Affiliation: UCB Pharma

Topic: Time-varying confounding/target trials

Abstract: The use of symptomatic treatment (ST) in randomized controlled trials (RCTs) can complicate the interpretation of treatment effects, especially when aiming to assess the "direct" effect of the investigative treatment. Traditional approaches often censor or omit data after ST initiation, which can introduce bias and inefficiency, particularly when ST use is common and its effect on outcome is confounded. G-estimation addresses this issue by "de-mediating" the effect of ST on outcomes. It achieves this through structural nested mean models (SNMMs), which quantify the causal effect of ST conditional on the history of time-varying confounders. This method retains all data, enabling precise estimates even with frequent ST use; in addition, it tends to outperform inverse probability of censoring weighted analyses which lack robustness when, as often, there are strong predictors of ST use.However, the uptake of g-estimation in the clinical trial context has arguably been limited. In order to stimulate further interest in this methodology and facilitate its use, particularly in the drug development space, we have created an illustrative 1-page guide that explains its essential elements, mathematically, visually, and with pseudo-R code snippets. Our guide provides a step-by-step guide to the estimation algorithm, including the calculation of visit-wise ST initiation probabilities (propensities) and the quantification of lagged ST effects. With this guide, it is our hope to highlight the potential of g-estimation to enhance the interpretation of RCT results by effectively handling complexities introduced by ST use.

Additional Authors: Stijn Vansteelandt

Abstract number: 58

Date: Wednesday 9th April

Title: Target trial emulation of dual treatment analyses: challenges of longitudinal TMLE

Presenter: Anna Menacher

Affiliation: Novo Nordisk

Topic: Time-varying confounding/target trials

Abstract: With more and more treatment options becoming available, methodological advancements for questions regarding the simultaneous use of multiple treatments will become increasingly more important. Here, we present a study on the effect of combination treatment with a GLP-1 receptor agonists and SGLT2 inhibitors, compared to older second-line antidiabetic drugs, such as dipeptidyl peptidase-4 inhibitors, sulfonylurea, or thiazolidinediones, on cardiovascular outcomes in people living with type 2 diabetes by emulating three target trials in UK electronic healthcare records. We utilise longitudinal targeted maximum likelihood estimation (LTMLE) to estimate the 2-year absolute risk differences and thereby account for time-varying confounders. With this work, we hope to highlight the challenges of longitudinal data analysis, such as near positivity violations, and event scarcity, and showcase the impact of various design choices of LTMLE, such as g-bounds, variance estimators, on-treatment definitions, superlearner algorithms, stratified vs pooled LTMLE, and other options that impact the estimation of absolute risks or their respective variance.

Additional Authors: Thomas Gerds, Kajsa Kvist

Abstract number: 59

Date: Wednesday 9th April

Title: Handling informative patient monitoring in routinely-collected data used to estimate treatment effects, with application to high-frequency hospital data

Presenter: Leah Pirondini

Affiliation: London School of Hygiene and Tropical Medicine

Topic: Time-varying confounding/target trials

Abstract: Introduction and Objectives:Routinely-collected hospital data provide opportunities to gain understanding of treatment effects that would not be feasible in randomised trials and that reflect their impact in realistic clinical practice. A challenge presented by hospital data is that measurements of patients' clinical status are made at high frequency, on differing schedules for each patient dependent on their underlying clinical status, so timing and frequency of measurements is informative. However, many existing causal inference methods assume measurements are made at regular time intervals. The aim of this work is to evaluate methods for estimating causal effects of longitudinal treatments in the presence of informative monitoring. This is motivated by hospital data on patients with COVID-19 and questions about optimal mechanical ventilation strategies. Methods and Results:We compare methods based on (i) marginal structural models fitted by inverse probability of treatment weighting (MSM-IPW), (ii) G-computation, and (iii) longitudinal targeted maximum likelihood estimation (LTMLE). We assume an underlying grid of time, such that time-dependent variables are either monitored or unmonitored at each time-point. Methods are based either on imputation of unmonitored covariate data (G-computation) or on adapting inverse probability weights to account for monitoring variables (MSM-IPW and LTMLE). We evaluate methods using a simulation study, comparing against more simple approaches using last-observation-carried-forward (LOCF) ignoring informativeness of monitoring. Data are simulated to represent a range of realistic scenarios with time-varying treatment and covariates, in which monitoring depends on past covariate, treatment and monitoring levels. We also illustrate methods in a real-world example using routinely-collected intensive care data from UCLH to investigate the use of invasive mechanical ventilation vs non-invasive or no ventilation on mortality in COVID-19 patients.We show that ignoring monitoring can result in bias, the size of which depends on informativeness of the monitoring process. All methods reduce bias compared with their naïve LOCF-based equivalents, with LTMLE and G-computation based methods resulting in the smallest bias. Conclusions:Data with informative monitoring are common in observational studies, but there is a lack of readily-implementable methods to handle them. We describe three methods and evaluate their performance.

Additional Authors: Karla Diaz-Ordaz, Ruth Keogh

Abstract number: 60

Date: Wednesday 9th April

Title: A comparison of statistical and computational tradeoffs in Sequential Target Trial Emulation approaches

Presenter: Carlos Poses

Affiliation: Real World Evidence Team, Department of Data Science and Biostatistics, UMC Utrecht

Topic: Time-varying confounding/target trials

Abstract: Target trial emulation is an increasingly popular framework for estimating the causal effects of interventions using (large-scale) electronic health records. Its main advantage is to transparently address both confounding and time-zero misalignment (often the culprit of immortal time and selection bias). Since treatment is typically initiated by exposed units on different calendar dates, researchers often emulate a sequence of trials with different time-zero points (Hernán, 2008). An open question in this setting is how to best utilize information from potential control units. One approach is to use all eligible controls at each initiation time and adjust for lack of exchangeability between groups using some confounder adjustment technique (e.g., Caniglia et al., 2021). An alternative approach is one-to-one matching (e.g., by matching both on time-zero and a set a confounders), effectively using only one control unit for each exposed unit (e.g., Barda et al. 2021). These approaches represent two extreme ends of a continuum of strategies, each with advantages and disadvantages. Concretely, while one-to-one matching is an intuitive, non-parametric confounding adjustment technique, it completely discards information from non-controls and restricts the estimands that can be studied. And while using all eligible controls does not discard any information and allows for other target estimands, it comes with its own challenges. First, increased computational costs may be non-trivial, especially for typically very large healthcare databases. Second, confounder adjustment often relies on parametric methods, and adding information from control units that are very different to the treated may increase, and not reduce, the variance of estimators (e.g., via very low or high weights if using inverse probability of treatment weights). In-between these two extremes, hybrid approaches, where some confounders are matched on, but remaining confounder imbalances are adjusted for via other (parametric) confounder adjustment technique, may offer a reasonable compromise. In this study, we investigate these two extremes as well as novel hybrid approaches in terms of bias, estimators' variance, computational costs and target estimands. We compare the theoretical properties of each approach, as well as present the design and preliminary results of a Montecarlo simulation with different data-generating conditions.

Additional Authors: Linda Nab, Miriam Sturkenboom, Oisín Ryan

Abstract number: 61

Date: Wednesday 9th April

Title: HOW DOES MODE OF BIRTH AFFECT THE RISK OF DEATH OR BRAIN INJURY IN VERY PRETERM BABIES? METHODOLOGICAL CONSIDERATIONS FOR A TARGET TRIAL EMULATION APPROACH

Presenter: Nicola Reeve

Affiliation: Cardiff University

Topic: Time-varying confounding/target trials

Abstract: Very preterm babies are at risk of poor neurodevelopmental outcomes and death. Intraventricular haemorrhage (IVH) after birth is the most prevalent of the poor outcomes. Birth by caesarean section (CS) may protect against IVH in very preterm babies, but evidence is limited. The ultimate aim of this study is to identify and obtain the quantitative evidence to inform a future definitive clinical trial to determine optimal mode of birth for preterm babies.For the analysis, we have access to routinely recorded, de-identified, national clinical data held in the National Neonatal Research Database; including all infants admitted to NHS Neonatal Units in England and Wales between 2012 and 2022, born between 22+0 and 31+6 weeks of gestation in singleton or twin births.The primary outcome is a composite of death or survival with severe IVH. We also investigate a range of secondary outcomes, aligned with the core outcome set for neonatal research.In this presentation, we discuss an Emulated Target Trial (ETT) approach to answer the question of how mode of birth (vaginal or CS) affects the outcomes of interest, focussing on some key methodological concerns. First, there is the important issue of what exactly the relevant clinical question is: as well as the effect of actual mode of birth, also key is the effect of intended mode of birth. The latter can (imperfectly) be deduced from the former together with information on whether or not induction of labour was attempted. This raises further issues around the timing of the intention, the subgroup of women for whom the question is relevant (and therefore the target trial's eligibility criteria), and subtle issues such as the interpretation of gestational age (an important confounder), which is itself (partially) affected by mode of birth (since delivery by CS often happens a few days faster than vaginal delivery, induced or otherwise). Estimation is via inverse probability weighting with confounders including gestational age, maternal comorbidities, parity and presentation.We discuss how the results of the ETT analyses can inform a future trial, in terms of eligibility criteria, estimands. sample size and potential challenges in recruitment.

Additional Authors: Jane Barnett, Judith Cutter, Rhian Daniel, Chris Gale, Dimitrios Siassakos, David Odd

Abstract number: 62

Date: Wednesday 9th April

Title: Target Trial Emulation for the effect of Dialysis-duration on the Survival Benefit of Kidney Transplant

Presenter: Dries Reynders

Affiliation: Department of Applied Mathematics, Computer Science and Statistics, Ghent University; Ghent (Belgium)

Topic: Time-varying confounding/target trials

Abstract: End Stage Renal Disease (ESRD) is the last phase of Chronic Kidney Disease, where the patients' kidneys no longer work sufficiently. In this stage, renal replacement therapy (RRT) is required, either dialysis or kidney transplant. With only a limited supply of donor kidneys, assessing their impact on survival is important and challenging. Numerous studies have advocated the advantage of receiving transplant without initiating dialysis first - but conclusive evidence is hard to find. In a setting unsuitable for RCTs, older observational studies often fail to address confounding-, selection-, lead time- and immortal time-bias. Recently, Target Trial Emulation (TTE) in many forms has been applied, leading to more substantiated evidence in favour of kidney transplant. What remains somewhat underexposed is how this specific setting impacts the assumptions one can reasonably make and the causal questions that may be relevant. A population-average treatment effect, for example, targets an effect in a world that is currently unattainable with the available donors. A per protocol analysis, contrasting transplant with never-transplant on the other hand, seems to ignore that 'never transplant' in reality is seldom a deliberate choice, but a result from the circumstances.We present results form a TTE-analysis on Stockholm-based ESRD patients in SCREAM with access to extensive and up-to-date repeated laboratory results and diagnoses in primary, secondary and tertiary care, allowing for thorough control of time varying confounding. We set up sequential trials at different times since the start of RRT comparing residual survival between patients transplanted then and counterparts continuing on dialysis at that time. A treatment x prior dialysis duration-interaction allows the subsequent transplant-benefit to depend on the waiting time on dialysis. While the conditional hazard ratio changes dramatically with the waiting time on dialysis, the residual 8-year survival difference is surprisingly invariant across the shrinking subpopulations of survivors. In the current allocation system, kidney transplant impact on residual survival does not significantly vary over recipient's prior time on dialysis. For patients, of course, any cost benefit based decision on when to transplant will additionally need to account for the chance of surviving up to a planned transplant time.

Additional Authors: Marie Evans ; Juan-Jesus Carrero ; Els Goetghebeur

Abstract number: 63

Date: Wednesday 9th April

Title: Emulating non-inferiority target trials: challenges and considerations

Presenter: Sabine Landau

Affiliation: King's College London

Topic: Time-varying confounding/target trials

Abstract: The target trial framework (TTF) is increasingly used to address causal questions about treatments using observational data. While most efforts have focused on emulating superiority trials (designed to determine whether one treatment is more effective than another or no treatment), there has been comparatively less focus on emulating non-inferiority trials. These trials are designed to show that a new treatment is not worse than a comparator by more than a pre-specified non-inferiority margin. For time-to-event outcomes, regulatory standards often require that the upper bound of the 95% confidence interval for the hazard ratio comparing the new treatment with the comparator is below a pre-specified non-inferiority margin, typically a value slightly greater than one. This has implications for how a target trial is emulated. Emulating a non-inferiority target trial differs from a superiority target trial in two ways. First, the estimand of interest must assess the causal effect of treatment receipt rather than treatment offer, as small offer effects may arise due to factors such as patient non-compliance or deviations from the intended treatment regimen. Second, for sustained treatments (e.g., repeated prescriptions), it is necessary to account for time-varying confounding in addition to baseline confounding to achieve unbiased estimation. In this work, we illustrate these principles by emulating the ROCKET AF trial (ClinicalTrials.gov ID: NCT00403767), a landmark cardiology non-inferiority trial, and provide insights into the practical challenges of applying the TTF to emulate non-inferiority target trials. We suggest that practitioners need to be aware that emulating a non-inferiority target trial is more complex than emulating a superiority target trial. Extra complexity arises due to the need to access more detailed observational information (e.g., time-coded information on treatment receipt and confounding variables) and the need to use more advanced causal inference methods (e.g., fitting marginal structural models for treatment regimens).

Additional Authors: Giulio Scola, Jack Wu, Nilesh Pareek

Abstract number: 64

Date: Thursday 10th April

Title: Improving interim decisions for single-arm trials by adjusting for baseline covariates and intermediate endpoints

Presenter: Eline Anslot

Affiliation: Ghent University

Topic: Applications in Health Sciences

Abstract: Researchers must endeavour to minimize exposure to harmful and/or ineffective treatments and maximize exposure to effective ones. One approach to achieve this is using multi-stage designs that allow stopping an ongoing trial for futility and/or efficacy. Commonly used designs for single-arm trials include group sequential designs with error spending functions, and the Simon two-stage design for binary endpoints. In both designs, the decision at the interim analysis may be improved, without making additional assumptions, by considering baseline covariates and/or intermediate endpoints.As participants usually enter the study at different times, some may not have reached their primary endpoint at the time of the interim analysis. However, they will possess other pertinent information such as baseline covariates and intermediate endpoints. Handling this as a missing data problem, prediction models are constructed to predict the primary endpoint of these subjects based on their available information. Our proposed interim estimator has the appealing property of being asymptotically unbiased, even when the prediction models are misspecified. When the prediction models are correct, the interim estimator is asymptotically efficient. A simulation study and data analysis were conducted to assess the operating characteristics of the proposed method. The results show a decrease in the chance of incorrectly stopping the trial for futility. This leads to an increase in power while controlling the type I error. Specifically, benefits are obtained with a) more partial information available relative to the total number of participants at interim, b) higher predictivity of the adjusted variable(s) and c) more correct prediction models.

Additional Authors: Kelly Van Lancker

Abstract number: 65

Date: Thursday 10th April

Title: Unveiling Sleep Dysregulation in Chronic Fatigue Syndrome (CFS) with and without Fibromyalgia (FM) through Bayesian Networks

Presenter: Michal Bechny

Affiliation: University of Bern

Topic: Applications in Health Sciences

Abstract: CFS and FM are two often co-occurring medically unexplained conditions associated with disrupted physiological regulation, including altered sleep patterns. Building on findings by Kishi (2011), who highlighted differences in sleep-stage (W, N1, N2, N3, REM) transitions, we developed a Dynamic Bayesian Network (DBN) to capture more detailed patterns of sleep and its dynamics. Female subjects (26 Healthy, 14 CFS, 12 CFS+FM) of homogeneous age (25-55 years) were carefully selected, with strict exclusion criteria for comorbidities, medications, blood tests, psychiatric symptoms, and other conditions, ensuring no hidden confounding and a focused investigation of CFS and FM effects. Previous research on non-clinical populations by Yetton (2018) suggested that sleep transitions follow a second-order Markov process, with improved prediction accuracy when incorporating stage duration. Using cross-validation and DBN of expertly predefined structures we systematically evaluated the impact of the incorporation of different Markovian lags (0-4), stage duration, time since sleep onset, and cumulative sleep duration. Our results confirmed the optimal lag = 2 (next-stage accuracy = 70.6%) and demonstrated that including stage durations significantly improved identifying conditioned subjects (AUROC = 75.4%). Time-since-onset and cumulative sleep did not enhance performance. Validation on independent observational datasets (Sleep Heart Health Study, Bern Sleep-Wake Registry) showed robust predictive performance, with accuracies of 60.1-69.8%, respectively. Further, by using mutilated Bayesian networks that fixed the optimal DBN for healthy, CFS, and CFS+FM groups, we simulated causal counterfactuals. CFS sleep showed significantly longer W, N3, and REM bouts, reduced REM prevalence, and altered transitions, such as increased (N1, REM)-to-W and decreased N1-to-REM and REM-to-N1. In CFS+FM, disruptions were more pronounced, with increased (W, N3) bouts, decreased (N1, N2) bouts, altered stage prevalences (increased N2, N3, reduced N3, REM), and significant disruptions in transitions, including increased (W, N1, REM)-to-N2, N3-to-(W, N1), and N2-to-N3, and decreased N3-to-N2, N1-to-REM, and (W, N2, REM)-to-N1. Lag-two transitions provided further insights into the distinct sleep dynamics of both conditions. Our study reveals unique disruptions in stage durations and transitions, enhancing the understanding of physiological mechanisms, supporting clinical differentiation, and informing the development of targeted, individualized treatments for CFS and CFS+FM.

Additional Authors: Francesca Faraci, Athina Tzovara, Julia van der Meer, Marco Scutari, Benjamin Natelson, Akifumi Kishi

Abstract number: 66

Date: Thursday 10th April

Title: kernel matching based on GPS in studying short-term effects of air pollution on population health

Presenter: Giulio Biscardi

Affiliation: Università degli Studi di Firenze

Topic: Applications in Health Sciences

Abstract: In studying the short-term effects of environmental pollutants on health outcomes (such as deaths and hospital admissions), the shape of the average dose-response function (aDRF) has important regulatory implications, and it is essential for calculating the health impact of exposures on the population. We frame the problem in the Potential Outcome (PO) approach of causal inference, defining POs for continuous exposures under SUTVA. Then, building on Wu et al. (2024), where GPS matching estimator is developed under local unconfoundedness and local overlap we propose an innovative method based on GPS kernel matching to estimate the aDRF. The key steps of our methods are the following. We first divide the exposure into intervals, e.g based on quantiles, and take the median point of each of them. We then calculate the GPS at each midpoint for each unit. For each interval the POs at the median point for all units are imputed using a weighted mean of the observed outcomes of the units belonging to that interval, where the weights are the product of two kernel distances, one between GPSs and the other between the exposures. Once the missing outcomes for each unit under each treatment level have been imputed, a flexible model for the outcome given the exposure values is specified. We conducted extensive simulation studies to evaluate the performance of the proposed method also with respect to existing competitors for estimating an aDRF. We apply the method to evaluate the short-term effects of PM2.5 on deaths from all causes in the metropolitan area of Florence (Italy, 2008-2020).

Additional Authors: Chiara Marzi, Alessandra Mattei, Michela Baccini

Abstract number: 67

Date: Thursday 10th April

Title: Assessing Principal Causal Effects with Outcome-dependent Sampling Using Principal Score Methods: An Application to the E3N Cohort

Presenter: Lisa Braito

Affiliation: University of Florence

Topic: Applications in Health Sciences

Abstract: Outcome-dependent sampling design, commonly known as case-control studies in epidemiology, are widely used to estimate the impact of an exposure on the (binary) outcome of interest. These designs involve sampling a group of individuals from a target population conditional on observed outcomes. Thus, outcome-dependent sample are inherently retrospective studies. While conventional methods such as logistic regression remain prevalent for analysing case-control data, they may be limited when it comes to drawing valid causal inferences. Furthermore, when the goal extends beyond assessing the average causal effect of an exposure on the outcome and aims at understanding the underlying causal mechanisms through intermediate variables, new methodological approaches are required. We show that directly applying principal stratification methods to case-control designs can lead to biased estimates and misleading conclusions. To address these issues, we propose a method for estimating principal causal effects in such designs, drawing on well-established principal score methods from the principal stratification literature in observational studies. Our approach leverages external auxiliary information about the population of interest to analyse data derived from outcome-dependent samples.The primary objective of this study is to investigate the relationship between menopausal hormone therapy (MHT), mammographic density, and breast cancer (BC) risk leveraging principal stratification approach. Indeed, scientific evidence suggests that MHT increases BC risk, and this increase may be partially mediated by mammographic density. The data comes from a nested case-control study within the French E3N prospective cohort study. To validate our methodology, we take advantage of a secondary intermediate variable, such as body mass index (BMI), which is measured across the entire cohort. By simulating case-control samples from the full cohort and comparing these results to those obtained from the entire cohort analysis, we are able to assess the robustness and accuracy of our proposed method. Our findings enhance causal inference methods for complex sampling designs, offering a robust framework for addressing epidemiological questions.

Additional Authors: Fabrizia Mealli, Vittorio Perduca, Gianluca Severi

Abstract number: 68

Date: Thursday 10th April

Title: External reproduction of a proxy-based causal model for estimating average effects of sequential vs concurrent chemo-radiotherapy on survival in stage III Non-Small Cell Lung Cancer (NSCLC)

Presenter: Charlie Cunniffe

Affiliation: The University of Manchester, Division of Cancer Sciences

Topic: Applications in Health Sciences

Abstract: PurposeRandomised controlled trials (RCTs) offer the most reliable evidence for treatment decisions. However, in cancer care, frail, elderly, and disadvantaged patients are often underrepresented, causing uncertainty about optimal treatment strategies, such as whether sequential or concurrent chemo-radiotherapy yields better outcomes. Observational causal inference may supplement RCT evidence; however, confounding factors that are not directly observed remain a challenge. Using available proxy measurements to infer these variables offers a potential solution. In the cancer setting, the patient's overall fitness is an important unobserved confounder for treatment and outcome, for which proxies like "performance score" are widely available within patient records. This study reproduces a recently introduced proxy-based causal inference method to assess transportability over populations with different data structures and treatment guidelines.MethodThis study employs proxy-based individual treatment effect modelling in cancer (PROTECT) to estimate the individual treatment effect of concurrent vs sequential chemo-radiotherapy on overall survival in 1117 routinely treated patients with stage III NSCLC seen between 2013 and 2023. A local model was developed by adapting the PROTECT directed acyclic graph to include our selected proxies of patient fitness – performance, comorbidity, and frailty scores. While the causal effect is generally not identifiable in the presence of unobserved confounding, PROTECT assumes that the proxies are conditionally independent given the unobserved confounder and incorporates domain knowledge via functional constraints on a Bayesian latent factor model. Individual treatment effect estimates for each patient are averaged to get the population's average treatment effect (ATE) estimate, reported as a hazard ratio. We compared the ATE to that from the PROTECT development study and the results of a meta-analysis of RCTs.ResultsThe ATE on survival in our population is 0.92 [0.73, 1.15] in favour of concurrent chemo-radiotherapy (the PROTECT development publication reports 1.01 [0.68,1.53]). The RCT meta-analysis reports a stronger effect size (0.84 [0.74,0.95]) than the causal analysis but has a younger (19% $>$ 70 vs 43%) and fitter (50% ECOG 0 vs 27%) population than our study. ConclusionRepeated analysis in two separate populations yielded comparable results, implying a robust estimation of the ATE. PROTECT is a promising new method that can be applied more broadly in cancer and other settings.

Additional Authors: Wouter van Amsterdam (University Medical Center Utrecht), Matthew Sperrin (The University of Manchester), Rajesh Ranganath (New York University), Fiona Blackhall (The University of Manchester), Gareth Price (The University of Manchester)

Abstract number: 69

Date: Thursday 10th April

Title: Integrating causal reasoning in academic curricula: exploring barriers and solutions

Presenter: Annick De Paepe

Affiliation: Ghent University

Topic: Applications in Health Sciences

Abstract: Introduction. Despite several calls for a more widespread use of causal inference methods in applied health research, their adoption remains limited. The integration of causal inference methods in academic curricula may be an important strategy to increase their use in future research, but explicit causal reasoning is only rarely included in introductory statistics or methodology courses. We aimed to gain insight into the barriers for integrating causal inference methods in academic curricula and to identify strategies to overcome these barriers. Methods. 26 researchers working across Europe with diverse backgrounds (statistics, health sciences, psychology) and with an interest and/or expertise in causal inference methods attended an in-person meeting on the application of causal inference methods in Psychology and Health Sciences. They were presented with two open-ended questions: (1) What are the main barriers to incorporate causal inference methods into statistics or methodology courses?, (2) What is the most important next step to advance teaching and research in this area? Participants submitted their responses anonymously, yielding 47 responses to the first and 61 responses to the second question. Responses were discussed within the group. Data were analysed using thematic analysis with codes based on behaviour change theories (Capability Opportunity Motivation - Behaviour (COM-B) model and the Behaviour Change Wheel (BCW)). Results. Barriers were identified in each of the categories of the COM-B model (psychological capability, Physical and Social Opportunity and Motivation). Examples of barriers include a lack of knowledge and limited experience using causal inference methods, limited number of hours of the statistical courses, lack of consolidation of causal inference methods across different courses, a lack of support from peers, a lack of self-efficacy to teach the subject and resistance by colleagues. Specific examples of intervention strategies and policies that could promote the integration of causal inference methods in introductory statistics and methodology courses were identified based on the BCW.Conclusion. Relevant barriers to integrate causal inference methods in statistics and methodology courses and strategies and policies to overcome these were identified. This knowledge could be used to develop interventions promoting the use of causal inference methods in teaching and future research.

Additional Authors: Jelle Van Cauwenberg, Tom Loeys, Stijn Vansteelandt, Kelly Van Lancker, Johan Steen, Geert Crombez, Oliver Dukes, Ruth Keogh, Camila Olarte Parra, Florian Schmiedek, Beatrijs Moerkerke, Els Goetghebeur, Thomas Matheve, Benedicte Deforche,

Delfien Van Dyck, Maïté Verloigne, Peter Tennant, Georgia Tomova, Jenny Van Beek, Vincent Brugger, Heidelinde Dehaene, Jeremy Labrecque, Dries Reynders, Ineke Van Gremberghe, Louise Poppe

Abstract number: 70

Date: Thursday 10th April

Title: Inference on sustained treatment strategies, with a case study on young women with breast cancer

Presenter: Elise Dumas

Affiliation: Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Topic: Applications in Health Sciences

Abstract: Lack of adherence can reduce the effectiveness of beneficial treatments. However, the extent to which adherence affects outcomes is unclear in many settings. For example, is it enough to adhere to treatment for 80% or 90% of the prescribed days? Does it matter whether adherence is higher earlier or later in the treatment schedule? And is it particularly important not to miss consecutive days of treatment? In this work, we explore a methodology to emulate and compare different adherence strategies. Specifically, we use a framework that incorporates explicit grace periods and regimes that depend on natural treatment patterns.Our work is motivated by a clinical question concerning women with early-stage hormone receptor-positive breast cancer, for whom daily endocrine therapy is prescribed for five to ten years. In these patients, young age is associated with both an increased risk of cancer recurrence and suboptimal adherence to endocrine therapy. Using French nationwide claims data, including more than 120,000 patients with breast cancer, we applied the proposed methods to compare the survival benefits achievable by implementing different adherence strategies to endocrine therapy for each age group. We emulated three different adherence strategies that allowed for gaps in treatment of no more than one, three, or six consecutive months. Our results show that young women would benefit substantially from stricter adherence to endocrine therapy, with treatment breaks never exceeding one month, highlighting the need for tailored strategies to improve treatment adherence in this population.

Additional Authors: Floriane Jochum, Florence Coussy, Anne-Sophie Hamy, Alena Majdling, Sophie Houzard, Christine Le Bihan-Benjamin, Fabien Reyal, Paul Gougis, Mats Julius Stensrud

Abstract number: 71

Date: Thursday 10th April

Title: Policy learning under constraint: seeking primary outcome maximization under adverse outcome control

Presenter: Laura Fuentes Vicente

Affiliation: Inria PreMeDICaL

Topic: Applications in Health Sciences

Abstract: Classically, a policy $\pi : \mathcal{X} \to \{0, 1\}$ maps patient characteristics $X \in \mathcal{X}$ to a treatment recommendation ($\pi(X) = 1$, `treat" versus $\pi(X)=0$,`do not treat"), in an attempt to maximize the objective function $t \mapsto \mathbb{E}(Y(t))$ with $Y(t)$ a real-valued, counterfactual primary outcome under intervention t. A policy may use the sign of the estimated Y-specific individual treatment effect (Y-ITE) to make recommendations. However, such an approach is ill-suited for multiple-outcome frameworks and disregards adverse events, which are critical when prescribing treatments in a clinical context. The challenge we tackle is to learn an optimal policy that maximizes such an objective function while satisfying constraints related to the real-valued, counterfactual adverse outcomes $Z(1), Z(0)$. For instance, the constraint can be the probability that $Z(\pi(X))$ fall within a satisfactory range, which ideally should be above a fixed, clinically determined threshold. We focus on situations where the adverse event is more likely when treatment is imposed, making the corresponding ITE consistently negative. The proposed approach uses both the sign of the estimated Y-ITE and the estimated probability that $Z(1)$ fall within a satisfactory range. We conduct a simulation study to assess the behavior of the approach from both an oracular and a statistical perspectives. We also assess their behaviors in a real-world in vitro fertilization (IVF) data scenario.

Additional Authors: Chambaz Antoine, Even Mathieu, Josse Julie

Abstract number: 72

Date: Thursday 10th April

Title: Effect modification in studies affected by "truncation by death"

Presenter: Bronner Gonçalves

Affiliation: University of Surrey

Topic: Applications in Health Sciences

Abstract: Epidemiologic studies in which outcomes are not defined for participants who die before end of follow-up are affected by "truncation by death", and crude analyses that condition on observed survival are not causal. In this context, the survivor average causal effect (SACE) has been commonly used as an estimand with a causal interpretation. However, due to differences in distribution of effect modifiers, the SACE may vary across study populations. In this work, we use the principal stratification framework to study the modification of the SACE by a variable that represents a possible common cause of survival and the outcome of interest, and by a variable that only affects survival. We show that the SACE can be expressed as a weighted average of this principal effect in each level of the effect modifier, and that weights depend on the frequency of the effect modifier in the "always-survivors" principal stratum. Although we focus on the risk difference measure, we also derive weights for the SACE defined as risk ratio. Finally, we discuss the implications of this work for the transportability of the SACE.

Additional Authors: Etsuji Suzuki

Abstract number: 73

Date: Thursday 10th April

Title: Refining the Allostatic Self-Efficacy Theory of Fatigue and Depression Using Causal Inference

Presenter: Alex Hess

Affiliation: ETH Zurich

Topic: Applications in Health Sciences

Abstract: Allostatic self-efficacy (ASE) represents a computational theory of fatigue and depression. In brief, it postulates that (i) fatigue is a feeling state triggered by a metacognitive diagnosis of loss of control over bodily states (persistently elevated interoceptive surprise); and that (ii) generalization of low self-efficacy beliefs beyond bodily control induces depression. Here, we converted ASE theory into a structural causal model (SCM). This allowed identification of empirically testable hypotheses regarding causal relationships between the variables of interest. Applying conditional independence tests to questionnaire data from healthy volunteers (N=60), we sought to identify contradictions to the proposed SCM. Moreover, we estimated two causal effects proposed by ASE theory using three different methods. Our analyses identified specific aspects of the proposed SCM that were inconsistent with the available data. This enabled formulation of an updated SCM that can be tested against future data. Second, we confirmed the predicted negative average causal effect from metacognition of allostatic control to fatigue across all three different methods of estimation. Our study represents an initial attempt to refine and formalize ASE theory using methods from causal inference. Our results confirm key predictions from ASE theory but also suggest revisions which require empirical verification in future studies.

Additional Authors: Dina von Werder, Olivia K. Harrison, Jakob Heinzle, Klaas Enno Stephan

Abstract number: 74

Date: Thursday 10th April

Title: The impact of dynamic and modified hemodynamic treatment policies on intraoperative hypotension: An application of causal inference methods in Anaesthesiology

Presenter: Markus Huber

Affiliation: Department of Anaesthesiology and Pain Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

Topic: Applications in Health Sciences

Abstract: Perioperative anaesthetic care involves the administration of medications to allow surgical interventions that would otherwise be intolerably painful. A challenging component of anaesthetic care is hemodynamic management by means of fluid administration to ensure adequate organ perfusion to avoid adverse events (e.g., stroke or acute kidney injury). Despite decades of research and clinical practice, optimal treatment regimes are unknown. Aditionally, the causal relationships underlying hemodynamic management, intraoperative blood pressure and clinical outcomes remain a crucial area to be investigated with causal inference methods.Here, we used a large public dataset (the INSPIRE dataset from Seoul, South Korea) to investigate the causal impact of different dynamic and modified hemodynamic treatment policies (DTP and MTP, respectively) on intraoperative hypotension (primary outcome). The primary outcome was defined as invasive mean arterial pressure (MAP) below 65 mmHg. The treatment policies were defined with respect to the main time-dependent covariate (this is, mean arterial pressure). The intervention was defined as the fluid administration of 250 mL plasma solution. For example, we evaluated the causal impact of administrating 250 mL of plasma solution only when intraoperative MAP dropped below 70 mmHg; otherwise, no plasma solution was administered (dynamic treatment policy). Another example is the modified treatment policy where an additional 250 mL of plasma solution was administrated (with respect to the observed, actual fluid administration) when intraoperative MAP dropped below 70 mmHg. To assess the clinical implications of the various treatment policies, we defined a time-dependent efficiency factor as the product of the incidence of hypotension, the percentage of patients treated with plasma solution and the amount of plasma solution administered. The treatment effect estimates were computed by Targeted Maximum Likelihood Estimation (TMLE) using the lmtp package.As results, we demonstrate that the various fluid treatment policies reduce intraoperative hypotension (incidence ranging from 20% to 35%) by about 5% when clinicians were to intervene only for low MAP thresholds (65 mmHg) to about 20% when clinicians were to follow a very pro-active treatment strategy (administering fluids already at high MAP thresholds of 75 to 80 mmHg). In particular, we found a long-term benefit of the proposed treatment policies, where the reduction in intraoperative hypotension increases over time. The analysis of the policies' efficiency suggested that interventions at low MAP

thresholds are the more efficient strategies.Overall, this study illustrated the opportunities for causal inference research in Anaesthesiology. We will further highlight challenges when applying causal inference methods in the analysis of perioperative data, for example irregular time intervals and distinct time scales between intra- and postoperative periods. We conclude by highlighting further areas of research in Anaesthesiology where causal inference is likely to play an important role in the future, notably in pain therapy and emergency medicine.

Additional Authors: Patrick Y. Wüthrich

Abstract number: 75

Date: Thursday 10th April

Title: Hybrid approaches, case studies in methods leveraging external data in randomized clinical trials

Presenter: Rima Izem

Affiliation: Novartis Pharma AG

Topic: Applications in Health Sciences

Abstract: This presentation will review methods and tools for prospectively planned hybrid approaches that incorporate clinical trial data with external data in the evaluation of treatment effectiveness and safety. First, the presentation will give an overview of existing design and analyses approaches leveraging external data in the pharmaceutical industry for internal decision making or for informing marketing authorization or reimbursement. Those designs include using historical data as external controls to complement single arm trials in oncology or rare disease indications, augmenting the control arm of a randomized controlled trial with external data, and using linkage to fill gaps in the journey of clinical trial participants with external sources. Then, the presentation will focus on analytical methods using causal inference methods with methods for integrating evidence from multiple sources such as hierarchical Bayesian models. Practical case studies will illustrate how a systematic identification of the target estimand and the evaluation of potential sources of bias can clarify the specifications of the main analyses and prioritize the quantitative bias analyses.

Additional Authors: Lisa Hampson

Abstract number: 76

Date: Thursday 10th April

Title: Effect of COVID-19 Vaccination Initiation on One-Year All-Cause Mortality Among Elderly Estonians: A Sequential Target Trial Emulation

Presenter: Anastassia Kolde

Affiliation: University of Tartu

Topic: Applications in Health Sciences

Abstract: When COVID-19 vaccines first became available, older individuals in Estonia were prioritized for immunization due to their higher risk of severe disease. However, Estonia faced notable vaccine hesitancy and had lower vaccination rates than the European Union average. While research supports the efficacy of COVID-19 vaccines in reducing disease-specific morbidity, their broader, longer-term effects on all-cause mortality among older populations remain unclear.This study aimed to evaluate the effect of initiating COVID-19 vaccination—regardless of vaccine brand— on one-year all-cause mortality among elderly Estonians who were prioritized during the early rollout phase of the vaccination campaign.We conducted a sequential target trial emulation using national healthcare and population registry data. Older individuals were included in the study upon becoming eligible for vaccination. Those who initiated vaccination were compared with those who did not, while taking into account potential confounding factors—such as comorbidities, healthcare utilization, region of residence, and educational attainment. The outcome of interest was one-year all-cause mortality starting from each individual's point of eligibility.Preliminary analyses suggest that early COVID-19 vaccination reduces one-year all-cause mortality among older Estonians, even in the context of substantial vaccine hesitancy and comparatively low vaccination coverage.By applying a sequential target trial emulation design, our study provides a stronger causal interpretation that timely COVID-19 vaccination reduces all-cause mortality among older, high-risk individuals. These findings underscore the importance of addressing vaccine hesitancy and ensuring accessible, prioritized vaccination to potentially improve long-term survival rates.

Additional Authors: Anneli Uusküla, Krista Fischer

Abstract number: 77

Date: Thursday 10th April

Title: Treatment effect estimation with counterfactual prediction using individual treatment plans: theory and application in radiotherapy

Presenter: Lotta M. Meijerink

Affiliation: Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

Topic: Applications in Health Sciences

Abstract: Background Within the context of radiotherapy, 'model-based clinical evaluation' was proposed, as an alternative approach to assess the benefit of new radiation technologies in terms of radiation-induced toxicities, when RCTs are not feasible or ethical. In this approach, the average treatment effect is estimated by comparing the actual observed outcomes under a new treatment technology with predicted outcomes for the same individuals if they had received the comparator technology. The approach takes advantage of a prediction model that uses not only patient characteristics as predictors, but also individual treatment details, such as the radiation dose to normal tissues from an individual treatment plan that an individual would have received under the comparator treatment technology.Objectives To provide methodological guidance to potential users of this approach, both within and outside of radiotherapy. Methods We used principles from causal inference to formalize the method and systematically identify the conditions necessary for its validity. Furthermore, we described sensitivity analyses and strategies to assess the plausibility of the conditions.To illustrate the approach, and to discuss the plausibility of the identifiability conditions, we applied the method to data of head- and neck cancer patients, to evaluate the potential benefit of proton therapy over photon therapy in reducing radiation-induced toxicity. Results Important conditions of the model-based clinical evaluation approach include model quality in the development population, and consistency of the individual treatment details. Another key condition is transportability of the outcome risk. We show, using the case study as illustration, that model calibration - under some circumstances - can provide evidence of complying with the transportability assumption, and that model recalibration can be used as a type of sensitivity analysis. Conclusions The model-based clinical evaluation approach holds potential in evaluating new (personalized) treatment strategies or technologies. However, as in every non-randomized study design, it comes with certain risks and relies on assumptions. Our guidance and illustrative case study provide researchers with a framework to assess the method's suitability and implement it effectively.

Additional Authors: Artuur M. Leeuwenberg, Jungyeon Choi, Johannes A. Langendijk, Judith van Loon, Remi A. Nout, Johannes B. Reitsma, Karel G.M. Moons, Ewoud Schuit

Abstract number: 78

Date: Thursday 10th April

Title: Use of Causal Inference in a Vaccine Development Program: A Case Study

Presenter: Joris Menten

Affiliation: Johnson & Johnson Innovative Medicine

Topic: Applications in Health Sciences

Abstract: IntroductionVaccine development programs are generally based on Randomized Controlled Clinical Trials which rely on randomization for causal conclusions. However, there is also a need for causal methods to address research questions that cannot fully be addressed using randomized comparisons. ApplicationIn this poster, we present two applications of the use of causal modeling in the development of a respiratory syncytial virus (RSV) vaccine for adults.(1) In vaccine development immune markers that help to accurately predict the efficacy of a vaccine are useful for regimen and dose selection and bridging between populations. These immunomarkers can be used as surrogate markers and are commonly referred to as Correlates of Protection (CoP). To establish the surrogacy value of a marker it is necessary to correlate the vaccine effect on the immune marker with the vaccine effect on the clinical outcome. Within a single trial, this requires information on the surrogate and/or outcome under both placebo and active vaccination. As each study participant receives either placebo or active vaccination, there is a need to use counterfactual outcomes to establish whether an immune marker can be labeled a CoP. Within the RSV vaccine development program, we used principal stratification to assess the surrogacy value of binding and neutralizing antibodies to RSV.(2) When using adeno-vectored vaccines, despite it was not observed in most of the programs using repeated adeno or adeno boosters, it is legitimate to raise the question of a potential reduced immunogenicity upon re-exposure. We explored the potential impact of vaccine induced anti-vector immunity in repeated vaccination with an adeno-based vaccine using two different inserts. We compared the immune response to an Ad26-based RSV vaccine in study participants who had received an Ad26-based COVID-19 vaccine 1-12 months prior to the RSV vaccination to those who did not receive an adeno-based vaccine. As prior exposure to the Adeno-based COVID-19 vaccine was not randomized, this comparison could be confounded by differences between exposed and unexposed participants in baseline characteristics that affect immune response. We used inverse probability weighting to correct for these differences. ConclusionsEven though clinical development programs tend to be based on clinical trials studies which use randomization to support causal conclusions, some research questions formulated in the program may require more complex causal inference methods.

Additional Authors: Cristina Sotto, Sjouke Vandenberghe, Mariska G.M. van Rosmalen

Abstract number: 79

Date: Thursday 10th April

Title: An examination by Principal Component Analysis of the effect of timing of food intake on BMI in the INRAN-SCAI 2005/06 nutrition survey

Presenter: Luigi Palla

Affiliation: Department of Public Health and Infectious Diseases, University of Rome La Sapienza

Topic: Applications in Health Sciences

Abstract: Introduction: Chrononutrition is an emerging field that studies the relationship between timing and regularity of food intake and their effect on health. Late main meals have been related to the risk of obesity, but main meals are often considered in isolation, rather than examining eating patterns throughout the day. In order to study both timing and regularity of eating, we used data from INRAN-SCAI 2005–06. Methods: A nationally representative cross-sectional sample was collected on 3323 subjects. The survey consisted of a 3-day diet diary recording every eating occasion, including the time and quantity of consumption. We derived the Diurnal Irregularity Eating Patterns (DIEPs) by Principal Component Analysis (with covariance matrix) jointly on average energy intake and irregularity of intake at 6 typical eating time intervals using the entire population. A linear regression model with robust standard errors to account for the correlation within household was applied including only adults (n=2313), with BMI as outcome, DIEP as exposure adjusting for covariates selected to standardize comparisons using inverse probability weighting (age, sex, civil status, profession, smoking, alcohol, breakfast intake, amount of weekly physical activity, geographic area, rural/urban environment, total daily energy intake, number of weekend days). The regression analysis was carried out including one of the 5 DIEP score in turn as exposure and also by dychotomising the score (based on the whole population and on adults only). Results: The first 5 DIEPs obtained for the 6 intervals explained 93% of variance and all proved interpretable: DIEP1 (var=47%) capturing energy intake at main meals, DIEP2 (var=22%) being a contrast between intakes at lunch and dinner, DIEP3 (var=10%) capturing snack time eating, DIEP4 (var=7.7%) breakfast eating, DIEP5 (irregular) night-eating. The crude estimates revealed a significant effect on BMI for most DIEPs but only DIEP4 showed significant effects across the adjusted models. Conclusion: Consistently with recent nutrition literature, the tendency to eat more calories at breakfast time (6-9AM) was associated with a lower BMI. Our approach included dimensionality reduction of chrononutritional exposures and control for likely confounding factors. Results are enhanced by the nationally representative nature of the sample but any causal interpretations is limited by the cross-sectional nature of the sample.

Additional Authors: Laura Lopez Sanchez; Bianca De Stavola

Abstract number: 80

Date: Thursday 10th April

Title: Thoughts after Droughts: Combined effects of prenatal and postnatal drought-related nutritional challenges on adolescent cognitive health in rural India

Presenter: Fabienne Pradella

Affiliation: Johannes Gutenberg University Mainz, Heidelberg Institute of Global Health, Stanford University

Topic: Applications in Health Sciences

Abstract: ObjectivesPoor maternal nutrition during pregnancy is associated with adverse health and human capital outcomes in the offspring over their life courses, yet the role of the postnatal environment for effect manifestation remains understudied. We investigate the interplay of pre- and postnatal drought-related nutritional shocks in shaping cognitive health among adolescents in rural India.MethodsMath and reading scores from the 2007-2018 Annual Status of Education Report (ASER) on 11-16-year-olds from rural India (N = 2,009,869) were linked to University of Delaware rainfall data. Rural India heavily depends on rainfed agriculture, so that rainfall variation proxies for nutritional shocks in a quasi-experimental set-up for causal inference. Drought was defined as rainfall below the 20th percentile of the district-specific long-term mean. We analyzed the interaction effect between prenatal drought exposure with drought exposure in childhood, while adjusting for, sex, age, year of assessment and district fixed effects.ResultsDrought exposure during pregnancy was associated with lower test scores in adolescence (reading, math, total score). Notably, children who experienced droughts in their second third life years were better prepared for drought circumstances if they had already been exposed to a drought prenatally, as indicated by positive interaction terms between prenatal and postnatal exposure. Effects did not differ by sex.ConclusionsPrenatal exposure to suboptimal nutrition might prepare the developing organism to deal with similar shocks in early postnatal life, with positive impacts on cognitive function years later. This finding is in accordance with epigenetic theories predicting that prenatal conditions sensitize developing organisms for similar postnatal circumstances.

Additional Authors: Sabine Gabrysch, Reyn van Ewijk

Abstract number: 81

Date: Thursday 10th April

Title: Data-Driven Causal Discovery in Lifestyle Modelling: Insights from a Longitudinal Study with Wearable Data

Presenter: Radoslava Svihrova

Affiliation: University of Bern

Topic: Applications in Health Sciences

Abstract: Over the past decade, mobile and wearable devices seamlessly integrated into people's daily lives, providing researchers with unparalleled opportunities to analyze lifestyle habits. Continuous monitoring, combined with user input and contextual data collected over extended periods, provides a foundation for exploring relationships between information derived from multi-source data. To identify these relationships, a causal Directed Acyclic Graph (DAG) is constructed in a data-driven way, incorporating the background knowledge about the temporal ordering of the measures and exogeneity of weather and calendar related variables. For the analysis features relevant for lifestyle modelling are crafted as aggregates on a daily level, and their lagged values are also considered. Given the real world setting, some missingness is expected in the data generating process, and hidden confounders may be present. To address these challenges and easily incorporate temporal ordering and exogeneity constraints, a Structural Expectation-Maximization (SEM) algorithm is employed. Further strengthening of the causal structure learning is done by providing an initial structure estimated in a data-driven way. For this the temporal Peter-Clark (tPC) algorithm with nonlinear independence tests is used as a computationally effective tool. Moreover, the robustness of the tPC is enhanced by a bootstrap. The resulting DAG is then evaluated for plausibility and compared against existing literature to ensure alignment with established findings or revealing new insights for further exploration. This approach highlights the feasibility of using wearable data for lifestyle insights, enabling the identification of meaningful patterns. The discovered structure will be used to inform further inference and serve as a starting point for the design of targeted lifestyle interventions. Future work will focus on extending the presented methodology with longer observation periods to uncover personalised lifestyle insights.

Additional Authors: Davide Marzorati, Alvise Dei Rossi, Tiziano Gerosa, Francesca Faraci

Abstract number: 82

Date: Thursday 10th April

Title: Synthetic control via covariate balancing with applications in long-term health impacts of extreme climate events

Presenter: Xiao Wu

Affiliation: Columbia University

Topic: Applications in Health Sciences

Abstract: Tropical cyclones pose a significant threat to the health and welfare of communities across the United States. While research has primarily focused on short-term impacts, typically within days or months of exposure, the long-term effects on the socioeconomic and demographic composition of communities remain underexplored. Understanding these long-term dynamics is critical for assessing community resilience and recovery. However, investigating multi-year health trends after tropical cyclones has been hindered by the limitations of existing statistical methods.Synthetic control methods have emerged as a promising tool for evaluating the long-term effects of public health interventions, but conventional approaches are not well-suited for large-scale environmental studies. Traditional methods are designed for small panel data and typically adjust for historical outcome trajectories and a limited number of covariates for a single exposed unit. In contrast, tropical cyclone studies require incorporating an extensive list of pre-exposure covariates for disaggregated spatial units. To address these challenges, we extend synthetic control methods by incorporating tailored balance conditions, formulated as a covariate balancing problem, which is solved using a convex optimization approach.Applying our method to data from all tropical cyclones in the United States between 2005 and 2018, we find that social vulnerability increased initially after exposure but decreased over the long term. One year post-cyclone, the social vulnerability index increased by 2.4% (95% CI: 0.7%–4.1%), persisting at elevated levels for at least five years. However, the social vulnerability index fell even lower than expected in the absence of a tropical cyclone, declining to -1.0% (95% CI: -1.9% to -0.0%) twelve years post-exposure. This phenomenon suggests the existence of long-term post-disaster gentrification, displacing original vulnerable communities and replacing them with newer, less vulnerable ones.

Additional Authors: Lingke Jiang, Yoshira Ornelas Van Horne, Robbie M Parks

Abstract number: 83

Date: Thursday 10th April

Title: "Explaining" the gender pay gap

Presenter: Christiane Didden

Affiliation: LMU Munich

Topic: Applications in Social Sciences

Abstract: The Oaxaca-Blinder (OB) method decomposes social disparities into two components: the explained component, related to group-specific differences in explanatory factors, and the unexplained component, related to group-specific differences in the associations of these factors with the outcome. This method can be linked to an interventional effects approach that simulates reductions in disparities by modifying explanatory factors, as shown by Jackson and VanderWeele (2018, Epidemiology). Building on their work, I explore the concept of "explained" in a decomposition analysis, using the gender pay gap as an empirical example and employing Directed Acyclic Graphs for illustration. The focus is on explanatory factors, X, that may lie on a backdoor path from exposure (gender) or mediator to outcome (wage), or on a front-door path from exposure to outcome, indicating that the explained component may result from either confounding or mediation. I categorize scientific questions into the three rungs of Pearl's causal ladder: • Rung 1 (Associational): E.g., "How strong is the association of X with gender and wage?" • Rung 2 (Interventional): "How would the gender pay gap change if X were manipulated?", e.g., "How would the gap change if the marginal distribution of X among women matched that of men?" • Rung 3 (Counterfactual): E.g., "Would women earn more if they had the values of X that they would have if they were men?"I connect these questions to the explained component, concluding that each rung has distinct implications for causal inference. In Rung 1, the "explained" component is purely descriptive, without causal interpretation. In Rung 2, the "explained" component reflects the causal effect of changing X on wage. In Rung 3, the "explained" component targets the natural (mediating) mechanisms through which exposure affects the outcome. However, Rung 3 involves counterfactuals that are empirically inaccessible and challenging to conceptualize when the exposure is ill-defined. I illustrate these points through an analysis of the gender pay gap using data from the German Socio-Economic Panel, with a focus on work experience, which I previously identified as a key factor for reducing the gap in an interventional effects analysis (Didden 2024, preprint on arXiv) .

Abstract number: 84

Date: Thursday 10th April

Title: Smoking heat Gun: estimating the effect of heatwaves on gun violence through a distance augmented synthetic control method

Presenter: Giulio Grossi

Affiliation: University of Florence

Topic: Applications in Social Sciences

Abstract: Gun violence in the United States has become an increasingly pressing issue in recent years. The upward trend is alarming, and policymakers are focusing their efforts on reducing the death toll. In this study, we examine gun violence in a changing world. Our primary focus is the relationship between heatwaves—direct consequences of climate change in intensity and frequency—and gun violence. We investigate these relationships within a potential outcomes framework, defining causal effects for areas that have experienced a heatwave and those potentially subject to spillover effects.We estimate the causal effect by introducing a spatially augmented version of the synthetic control method, leveraging the spatial information in the data to improve the interpretability of our estimates and reduce their variability. We employ a Bayesian regression approach to penalize the selection of more distant control units, within a semiparametric framework that balances unobserved spatial confounding.Our findings contribute substantively by clarifying how environmental factors relate to gun violence, and methodologically by naturally extending the synthetic control method to spatial data.

Additional Authors: Falco J. Bargagli Stoffi, Leo Vanciu

Abstract number: 85

Date: Thursday 10th April

Title: Assessing the Causal Impact of Early Tracking Postponement on Inequality of Opportunity: Evidence from the Italian Single Middle School Reform

Presenter: Kevin Taglialatela Scafati

Affiliation: University of Florence

Topic: Applications in Social Sciences

Abstract: Over the past two decades a growing body of empirical research has sought to evaluate the causal impact of school tracking on inequality. This paper contributes to this literature by exploiting the case of the 1963 Italian Single Middle School reform to investigate the causal effect of an early tracking postponement on inequality of opportunity (IOp) among students. Our primary outcome of interest is a long-term well-being measure, namely an estimate of the individual permanent income. By exploiting the reform's innovations and their imple-mentation timeline, we identify two groups of students exposed to educational systems that differ solely due to the presence of early tracking. These groups overlap in terms of key ascribed characteristics—such as birthplace, parental education, occupation, and sex—which serve as potential sources of inequality. Building on established social science literature, we employ inequality indices(e.g., Gini, MLD) applied to predicted individual well-being based on ascribed factors as measures of IOp. Our primary estimands contrast these measures across the two selected exposure groups, providing insights into the causal effect of interest. The nature of these novel estimands presents challenges for estimation; we adopt a Bayesian approach to inference which allows to naturally quantify the uncertainty of the targeted quantities and to flexibly specify the outcome model through a non-parametric BART specification.

Additional Authors: Paolo Brunori, Fabrizia Mealli

Abstract number: 86

Date: Thursday 10th April

Title: A Nonparametric Bayesian Approach for High-Dimensional Causal Effect Estimation in Survival Analysis

Presenter: Tijn Jacobs

Affiliation: Vrije Universiteit Amsterdam

Topic: Bayesian Causal inference

Abstract: Estimation of causal effects on survival in the presence of many confounders is hampered by the high-dimensionality of the data. Published work often addresses either high-dimensionality or survival data, but rarely both. We contribute to this gap by studying high-dimensional causal inference methods that account for confounding bias and the complexities introduced by censored observations. We present a nonparametric Bayesian method to estimate causal effects in high-dimensional, right-censored, and interval-censored survival data. Our approach employs a Bayesian ensemble of Additive Regression Trees (BART) with global-local shrinkage priors on the leaf node parameters (i.e., step heights). This contrasts with existing methods that induce sparsity through the tree skeletons. By focusing shrinkage on the leaf parameters, our method retains all covariates in the model. This reduces the risk of omitting relevant confounders and ensures compliance with the unconfoundedness assumption. The ensemble of Bayesian trees also captures complex, nonlinear relationships between covariates, treatment, and survival outcomes. We use an efficient Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm to sample from the posterior distribution of both the regression trees and the causal treatment effect. Our framework supports a wide range of global-local shrinkage priors. We demonstrate the performance of our method using the Horseshoe prior, which adapts to various levels of sparsity. The general implementation accommodates any scale mixture prior, providing a fast and flexible computational approach for high-dimensional data. We evaluate our method across a diverse set of simulated data settings, both sparse and dense. This evaluation showcases the method's robustness and flexibility. In sparse settings, the method effectively identifies confounders, while in dense settings, it captures intricate interactions between covariates, treatment, and outcomes. We demonstrate the practical utility of our approach on real-life data from cancer patients.

Additional Authors: Stéphanie van der Pas, Wessel van Wieringen

Abstract number: 87

Date: Thursday 10th April

Title: Population average treatment effect estimation under positivity violations: Adapting BART+SPL for high-dimensional covariates

Presenter: Lennard Massmann

Affiliation: University of Duisburg-Essen; Ruhr Graduate School in Economics

Topic: Bayesian Causal inference

Abstract: The positivity assumption is a fundamental requirement for causal inference in the potential outcomes framework, ensuring that all individuals have a positive probability of receiving each treatment option. However, real-world datasets often violate this assumption, particularly in regions of non-overlap, where one treatment group isunderrepresented or entirely absent for certain combinations of confounding variables. Traditional approaches, such as trimming and weighting, address these violations but typically modify the target population, potentially introducing bias. The Bayesian Additive Regression Trees with Spline Models (BART+SPL) approach has been proposed as a solution to this issue. BART+SPL combines BayesianAdditive Regression Trees (BART) for imputation in regions of treatment overlap with spline models (SPL) for extrapolation into non-overlap regions, preservingthe initial target population. This two-stage methodology reduces model dependence and improves uncertainty quantification in regions of non-overlap. However,BART+SPL's performance is compromised when dealing with high-dimensional covariates.To address this limitation, this paper proposes SBART+SPL, an extension of the BART+SPL framework that integrates SoftBART to enhance both precisionand uncertainty quantification when estimating population average treatment effects (PATE) in the presence of high-dimensional covariates and violations of the positivity assumption. A simulation study demonstrates that SBART+SPL yields better precision and improved coverage for both binary and continuous outcomes under these conditions. Additionally, the practical applicability of SBART+SPL is illustrated by reanalyzing two empirical case studies: the first evaluates the impact of exposure tonatural gas compressor stations on cancer mortality rates across U.S. counties, and the second investigates the effect of right heart catheterization on survival outcomesfor critically ill female patients.

Abstract number: 88

Date: Thursday 10th April

Title: Type 2 Tobit Sample Selection Models with Bayesian Additive Regression Trees

Presenter: Eoghan O'Neill

Affiliation: Erasmus University Rotterdam

Topic: Bayesian Causal inference

Abstract: This paper introduces Type 2 Tobit Bayesian Additive Regression Trees (TOBART-2). BART can produce accurate individual-specific treatment effect estimates. However, in practice estimates are often biased by sample selection. We extend the Type 2 Tobit sample selection model to account for nonlinearities and model uncertainty by including sums of trees in the selection and outcome equations. A Dirichlet Process Mixture distribution for the error terms allows for departure from the assumption of bivariate normally distributed errors. Soft trees and a Dirichlet prior on splitting probabilities improve modelling of smooth and sparse data generating processes. We include a simulation study and an application to the RAND Health Insurance Experiment data set.We also describe an alternative TOBART-2 implementation that marginalizes out both the terminal node parameters and the covariance term determining sample selection when updating trees in the outcome equation, and then makes a joint draw of all marginalized parameters. This is motivated by the observation of van Hasselt (2011) in the linear model setting that there is high dependence between the outcome equation coefficient draws and the covariance parameter draws.In addition to TOBART-2, this paper contains the following:To allow for data-informed calibration of the prior variance of the outcome equation errors, an existing prior is generalized with an additional data-informed hyperparameter. This is motivated by the fact that good performance of standard BART without hyperparameter tuning is partly attributable to well-calibrated hyperparameters. The sampler of Ding (2014) is improved by removing unnecessary imputation of the missing outcomes.We provide new insights pertaining to the van Hasselt (2011) and Omori (2007) Tobit-2 priors. The CDF of the correlation between selection and outcome errors is derived for both priors. We derive the conditions in which van Hasselt's (2011) implied prior on the correlation is unimodal, and when it is concentrated at +1 and -1. The prior CDF of the outcome variance is derived for prior calibration. It is observed that the prior correlation distribution specified by Omori (2007) can take various unimodal and bimodal forms, depending in non-trivially on hyperparameters. Options for prior calibration are discussed.

Abstract number: 89

Date: Thursday 10th April

Title: Bayesian estimation of heterogenous treatment effects from panel data

Presenter: Pantelis Samartsidis

Affiliation: University of Cambridge

Topic: Bayesian Causal inference

Abstract: Assessing the effect of an intervention (or policy/treatment) on a set of outcomes using observational time-series data is a problem that arises frequently in various fields, including public health and epidemiology. Often, intervention effects exhibit great heterogeneity across sampling units (e.g. hospitals or geographical regions) and over time. In such cases, average treatments effects are not necessarily informative, and we are instead interested in alternative causal estimands, mainly individual treatment effects (ITEs), which are common in the synthetic controls (SC) and generalised SC literature, or conditional average treatment effects (CATEs).In this work, we will discuss some important challenges arising when estimating ITEs and CATEs in such settings. First, the presence of unobserved confounders of which the total number may not be known in advance, and which need to be accounted for when defining and estimating CATEs. Second, the possibility that the intervention may affect the outcomes through various routes (e.g. vaccination reducing mortality both by boosting immune response and by preventing transmission), which leads to the notion of separable ITEs and CATEs. Third, the possibility that outcomes are of count type which complicates the estimation of ITEs under the standard assumption that potential outcomes are perfectly correlated after conditioning on confounders. Fourth, the difficulty in estimating the parameters of a model for treatment assignment when few units are treated.To address these challenges, we propose a Bayesian factor analysis model. Our model includes sub-models for treatment-free outcomes, outcomes under intervention, and the treatment assignment. A shrinkage prior is used to account for uncertainty in the total number of unobserved confounders, and complexity-penalising priors are used to reflect the prior expectation that few of the covariates modify the intervention effects. We then demonstrate how the issue of estimating the parameters of a treatment assignment model can be circumvented using a cut posterior approach, the properties of which are studied through extensive simulation studies. Finally, we develop copula-based approach to facilitate estimation of ITEs for count outcomes.We use our methodology to evaluate the impact of an intervention aiming to improve the effectiveness of contact tracing for COVID-19.

Additional Authors: Shaun Seaman, Robert Goudie, Lorenz Wernisch, Daniela De Angelis

Abstract number: 90

Date: Thursday 10th April

Title: Personalised Causal Effect Estimation with Observational and Experimental Data

Presenter: Sofia Triantafillou

Affiliation: University of Crete

Topic: Bayesian Causal inference

Abstract: A frequent goal in healthcare is to estimate personalised causal effects in order to select the best treatment for a patient from observational or experimental (RCT) data (or both), where "best" is defined in terms of optimizing a desired outcome. The first task in estimating personalized effects is selecting the optimal set of personalization covariates (causal feature selection). This set of covariates is the Markov Boundary of the outcome in the experimental distribution, also known as the Interventional Markov Boundary (IMB), and can be identified from RCT data using methods for finding Markov Boundaries. However, most RCT data are very limited in sample size, and do not work well with these methods. We develop methods that combine limited experimental and large observational data to identify the IMB, and improve the estimation of conditional (personalized) causal effects. The methods are based on Bayesian regression models. In simulated and semi-synthetic data, we show that our methods identify the correct IMB and improve causal effect estimation.

Additional Authors: Konstantina Lelova, Gregory F. Cooper

Abstract number: 91

Date: Thursday 10th April

Title: Causal Inference on Quantiles in High Dimensions: A Bayesian Approach

Presenter: Duong Trinh

Affiliation: University of Glasgow

Topic: Bayesian Causal inference

Abstract: This paper proposes a novel approach, Bayesian Analog of Doubly Robust (BADR) estimation, to estimate unconditional Quantile Treatment Effects (QTEs) in observational studies. By augmenting the proposed estimator with shrinkage priors, this framework can account for high-dimensional covariates and feature a flexible Bayesian modeling strategy with favorable frequentist properties in finite samples, even when either the treatment assignment or outcome models are misspecified. The proposed approach offers a straightforward and adaptable implementation for incorporating probabilistic machine learning techniques to fit the propensity score and conditional cumulative distribution function, followed by combining posterior draws. This enables the effective handling of high-dimensional covariate spaces or nonlinear relationships to achieve better accuracy and appropriate uncertainty quantification. The simulation results show that BADR estimators yield a substantial improvement in bias reduction for QTE estimates compared with popular alternative estimators found in the literature. We revisit the role of microcredit expansion and loan access on Moroccan household outcomes, demonstrating how the new method adds value in characterizing heterogeneous distributional impacts on outcomes and detecting changes in overall economic inequality, which is also appealing to other applied contexts.

Additional Authors: No

Abstract number: 92

Date: Thursday 10th April

Title: Causally Sound Priors for Binary Experiments

Presenter: Nicholas Irons

Affiliation: University of Oxford

Topic: Bayesian Causal inference

Abstract: We introduce the BREASE framework for the Bayesian analysis of randomized controlled trials with a binary treatment and a binary outcome. Approaching the problem from a causal inference perspective, we propose parameterizing the likelihood in terms of the baselinerisk, efficacy, and adverse side effects of the treatment, along with a flexible, yet intuitive and tractable jointly independent beta prior distribution on these parameters, which we show to be a generalization of the Dirichlet prior for the joint distribution of potential outcomes. Our approach has a number of desirable characteristics when compared to current mainstream alternatives: (i) it naturally induces prior dependence between expected outcomes in the treatment and control groups; (ii) as the baseline risk, efficacy and risk of adverse side effects are quantities commonly present in the clinicians' vocabulary, the hyperparameters of the prior are directly interpretable, thus facilitating the elicitation of prior knowledge and sensitivity analysis; and (iii) we provide analytical formulae for the marginal likelihood, Bayes factor, and other posterior quantities, as well as an exact posterior sampling algorithm and an accurate and fast data-augmented Gibbs sampler in cases where traditional MCMC fails. Empirical examples demonstrate the utility of our methods for estimation, hypothesis testing, and sensitivity analysis of treatment effects.

Additional Authors: Carlos Cinelli

Abstract number: 93

Date: Thursday 10th April

Title: Semiparametric Triple Difference Estimators

Presenter: Sina Akbari

Affiliation: EPFL

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: The triple difference causal inference framework is an extension of the widely known difference-in-differences framework. It relaxes the parallel trends assumption of the difference-in-differences framework through leveraging data from an auxiliary domain. Despite being commonly applied in empirical research, the triple difference framework has received relatively limited attention in the statistics literature. Specifically, investigating the intricacies of identification and thedesign of robust and efficient estimators for this framework has remained largely unexplored. This work aims to address these gaps in the literature. From an identification standpoint, we present outcome regression and weighting methods to identify the average treatment effect on the treated in both panel data and repeated cross-section settings. For the latter, we relax thecommonly made assumption of time-invariant covariates. From an estimation perspective, we consider a range of semiparametric estimators for the triple difference framework in both panel data and repeated cross sections settings. These estimators are based upon the cross-fitting technique, and flexible machine learning tools can be used to estimate the nuisance components. We demonstrate that our proposed estimators are doubly robust, and we characterize the conditions under which they are consistent and asymptotically normal. We evaluate our proposed estimators in extensive simulation studies.

Additional Authors: AmirEmad Ghassami, Negar Kiyavash

Abstract number: 94

Date: Thursday 10th April

Title: Causal Effects of Time-Varying Exposures: A Comparison of Structural Equation Modeling and Marginal Structural Models in Cross-Lagged Panel Research

Presenter: Jeroen Mulder

Affiliation: Universiteit Utrecht

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: A common question shared across research disciplines is how a time-varying exposure has a causal effect on a(n) (end-of-study) outcome. This question is often studied using panel data, in which the same people are observed multiple times on the same variables. In psychology and related fields, a popular modeling approach for such data is the use of various cross-lagged panel models within the structural equation modeling (SEM) framework. However, the use of SEM for causal inference is critiqued in some of the causal inference literature for unnecessarily relying on a large number of parametric assumptions, and alternative methods originating from the potential outcomes framework have been recommended, such as inverse probability weighting (IPW) estimation of marginal structural models (MSMs).Obviously, this claim should raise concerns among SEM users. Yet, disciplinary differences can hinder SEM-users to appreciate the arguments, concerns, and solutions put forward by SEM critics who come from fields like epidemiology and biostatistics. For example, in biomedical fields, the focus is commonly on a binary causal variable (the "treatment"), and when the state of this variable can vary over time, the focus is often on contrasting an end-of-study outcome following two different treatment regimes (i.e., joint effects). In contrast, in psychology the focus is typically on bidirectional lag-1 effects between time-varying X and Y variables. Yet, literature comparing these different estimands and estimation techniques from different disciplines is scarce. To better understand this criticism on the use of SEM, we describe three phases of causal research. Using an empirical example, we explain (differences in) the assumptions that are made throughout these phases for a SEM approach versus IPW regression of linear MSMs (IPW-MSM). Moreover, using simulations we compare the finite sample performance of a SEM approach and IPW-MSM for the estimation of time-varying exposure effects on an end-of study outcome under violations of parametric assumptions. Our results nuance some of the criticism of SEM, but we emphasize that these results should certainly not be interpreted as an incentive to continue currently popular SEM modeling practices, when the actual goal is causal inference.

Additional Authors: Kim Luijken, Bas B. L. Penning de Vries, Ellen L. Hamaker

Abstract number: 95

Date: Thursday 10th April

Title: The Impact of COVID-19 Pandemic on Non-deferrable Diseases

Presenter: Sara Muzzì

Affiliation: University of Milan-Bicocca

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: The COVID-19 pandemic has profoundly disrupted healthcare systems globally, raising concerns about its impact on non-COVID-19 patients requiring immediate and intensive cares. This paper investigates the effects of the pandemic on the quality of care for Acute Myocardial Infarction (AMI) and Stroke patients in Lombardy, Italy. Taking advantage of rich administrative data (i.e. hospital discharges data, emergency call and mortality registries) and leveraging the national lockdown as an exogenous shock in a quasi-experimental framework, we estimate the causal effects of COVID-19 on in-hospital and out-of-hospital mortality rates, on changes in ambulance response time and discussing their implications for the regional health care system. Our results reveal a 60% increase in daily out-of-hospital deaths during the pandemic and significant delays in ambulance response times, with an average increase of 11 minutes for both AMI and Stroke patients. However, in-hospital mortality remained stable, suggesting that delays in ambulance transport did not directly affect outcomes for patients who reached hospitals.

Additional Authors: Paolo Berta and Stefano Verzillo

Abstract number: 96

Date: Thursday 10th April

Title: A joint test of unconfoundedness and common trends

Presenter: Eva Oeß

Affiliation: Universität zu Köln

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: This paper introduces an overidentification test of two alternative assumptions to identify the average treatment effect on the treated in a two-period panel data setting: unconfoundedness and common trends. Under the unconfoundedness assumption, treatment assignment and post-treatment outcomes are independent, conditional on control variables and pre-treatment outcomes, which motivates including pre-treatment outcomes in the set of controls. Conversely, under the common trends assumption, the trend and the treatment assignment are independent, conditional on control variables. This motivates employing a Difference-in-Differences (DiD) approach by comparing the differences between pre- and post-treatment outcomes of the treatment and control group. Given the non-nested nature of these assumptions and their often ambiguous plausibility in empirical settings, we propose a joint test using a doubly robust statistic that can be combined with machine learning to control for observed confounders in a data-driven manner. We discuss various causal models that imply the satisfaction of either common trends, unconfoundedness, or both assumptions jointly, and we investigate the finite sample properties of our test through a simulation study. Additionally, we apply the proposed method to five empirical examples using publicly available datasets and find the test to reject the null hypothesis in two cases.

Additional Authors: Martin Huber

Abstract number: 97

Date: Thursday 10th April

Title: Conditional Independence Testing in Time Series

Presenter: Jieru Shi

Affiliation: University of Cambridge

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: Granger causality has traditionally been studied under the assumption of a linear vector autoregressive (VAR) model, with tests focusing on the significance of the VAR coefficients. We address the problem of testing a model-free null hypothesis of conditional independence in time series—specifically, whether Y_{t+1} and X_t are conditionally independent given the history of Y up to time t. We propose nonlinearly regressing both of them on the history of Y up to time t, and calculating a test statistic based on the sample covariance of residuals, called the Generalized Temporal Covariance Measure (GTCM). The type I error control of the test relies on the relatively weak assumption that user-chosen regression procedures estimate conditional means at a sufficiently fast rate that is slow enough to accommodate nonparametric settings. By further assuming stability of the regression procedures and weak dependence in the time series, we can utilize the entire dataset to estimate the conditional means without splitting the time series into subsets.

Additional Authors: Rajen D. Shah

Abstract number: 98

Date: Thursday 10th April

Title: Difference-in-Differences with Non-Ignorable Attrition

Presenter: Javier Viviens

Affiliation: European University Institute

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: Unbalanced panels are frequently used in Difference-in-Differences (DiD) applications. In this paper, I employ principal stratification analysis to highlight the potential drawbacks of the DiD research design when the outcome is missing for some units. Specifically, the conventional ATT estimand may not be well defined, and the DiD estimand cannot be interpreted causally without additional assumptions. To address these issues, I develop an identification strategy to partially identify causal effects on the set of units for which the outcome is observed and well-defined under both treatment regimes. I adapt Lee bounds to the DiD setting, replacing the unconfoundedness assumption in the original trimming strategy proposed by Lee (2009) with a principal parallel trend assumption. I also explore how to leverage multiple sources of attrition to relax the monotonicity assumption, thereby allowing the four latent strata to exist, which may be of independent interest. Alongside the identification results, I present estimators and their asymptotic distributions. I illustrate the relevance of the proposed methodology by analyzing a job training program in Colombia.

Additional Authors: None

Abstract number: 99

Date: Thursday 10th April

Title: Enhancing Generalizability in Regression Discontinuity Designs: A Graphical Approach to Combining Multiple Fuzzy RDDs

Presenter: Julia Kowalska

Affiliation: Vrije Universiteit Amsterdam

Topic: Difference in differences / panel data / regresssion discontinuity

Abstract: Regression discontinuity design (RDD) provides robust estimates of the local treatment effect at the cutoff, leveraging that units close to the cutoff are as if randomized. However, their generalizability beyond the cutoff is limited. We explore the potential of combining data from multiple fuzzy RDDs that share the same score and outcome variables. Our work is motivated by the Dutch Arthroplasty Register dataset, which contains information on the primary total hip replacement from all Dutch hospitals. Specifically, we examine cases where hospitals used age-based cutoff criteria to decide the fixation type. As the cutoff points vary across the hospitals, it gives the opportunity to estimate the treatment effect for a broader population. We apply propensity score methods to meaningfully integrate the data, accounting for selection bias in hospital admissions. Moreover, for fuzzy RDDs, compliance behavior— categorized as always-takers, never-takers, and compliers— further complicates generalizability, as the treatment effect is identifiable only for compliers. Since compliance type is a latent variable often correlated with population characteristics, combining data from multiple fuzzy RDDs requires careful interpretation of the compliance types and their implications for estimating the treatment effect. To address this issue, we take a novel graphical approach, which has not yet been exploited in the context of the regression discontinuity design. Based on this approach, we formulate necessary assumptions and prove the identification of the average treatment effect with respect to the score value. Our framework provides a pathway to enhance the generalizability of RDD treatment effects while addressing challenges related to heterogeneous populations and latent compliance behaviors.

Additional Authors: Stéphanie van der Pas, Mark van de Wiel

Abstract number: 100

Date: Thursday 10th April

Title: Differential Network inference based on causal discovery

Presenter: Emilie Devijver

Affiliation: CNRS

Topic: Graphical models

Abstract: Comparing causal relationships across populations is essential in various fields. We addresse the inference of a differential network via causal discovery by proposing a novel method that operates without semi-parametric assumptions and handles continuous, discrete, and mixed data. We assume that the two different population arise from interventions on an unknown structural causal model, and we introduce the concept of interventional faithfulness. Our approach includes theoretical guarantees and demonstrates strong performance on simulated and real datasets.

Additional Authors: Daria Bystrova

Abstract number: 101

Date: Thursday 10th April

Title: Weak equivalence of causal graphs

Presenter: Søren Wengel Mogensen

Affiliation: Copenhagen Business School

Topic: Graphical models

Abstract: Different causal graphs may encode the same set of conditional independences in which case they are said to be Markov equivalent. Causal discovery algorithms often use tests of conditional independence to learn a Markov equivalence class of graphs. These tests have low power when using large conditioning sets. This led to the development of the anytime FCI algorithm which avoids using independence tests that condition on a large set of variables.We discuss the idea of weak equivalence which is a generalization of Markov equivalence. Two directed acyclic graphs with node set V are Markov equivalent if they agree on whether A and B are d-separated given C for all triples (A,B,C) such that A,B, and C are subsets of V. On the other hand, two causal graphs are said to be weakly equivalent if they encode the same set of conditional independences when restricting to some subset of all possible triples. We give different examples of relevant weak equivalence relations, and our main example is that of k-weak equivalence: Two graphs are k-weakly equivalent if they agree on all triples (A,B,C) such that the cardinality of C is less than or equal to k.In this talk, we present a new characterization of weak equivalence of directed acyclic graphs. This characterization is useful for causal discovery of weak equivalence classes, and we discuss how to apply this result. In particular, by characterizing k-weak equivalence we describe what can be learned by a causal discovery algorithm when only using conditional independence tests with small conditioning sets. This helps pave the way for feasible causal structure learning.

Abstract number: 102

Date: Thursday 10th April

Title: The Case for Time in Causal DAGs

Presenter: Alexander Reisach

Affiliation: Université Paris Cité

Topic: Graphical models

Abstract: We make the case for including time explicitly in the definition and interpretation of causal directed acyclic graphs (DAGs). Causality requires temporal precedence of the cause, meaning that a causal effect that exists between variables in one time order may not exist in another. Therefore, any causal model requires temporal qualification. We argue that an explicit treatment of time resolves existing ambiguity in causal DAGs and is essential to assessing the validity of the acyclicity assumption. If variables are separated in time, their causal relationship is necessarily acyclic. Otherwise, the absence of possible cyclic effects is an additionally requirement. We introduce a formal distinction between these two types of acyclicity and lay out their respective implications. We outline connections of our contribution with different strands of the broader causality literature and discuss the ramifications of considering time for the interpretation and applicability of DAGs as causal models.

Additional Authors: Antoine Chambaz, Sebastian Weichwald, Alberto Suarez

Abstract number: 103

Date: Thursday 10th April

Title: Revisiting Model Evaluation in Treatment Effect Estimation: A Critical Appraisal

Presenter: Hugo Gobato Souto

Affiliation: University of Sao Paulo

Topic: Heterogeneous treatment effects

Abstract: This paper critically examines current methodologies for evaluating models in Conditional and Average Treatment Effect (CATE/ATE) estimation, identifying several key pitfalls in existing practices. The current approach of over-reliance on specific metrics and empirical means and lack of statistical tests necessitates a more rigorous evaluation approach. We propose an automated algorithm for selecting appropriate statistical tests, addressing the trade-offs and assumptions inherent in these tests. Additionally, we emphasize the importance of reporting empirical standard deviations alongside performance metrics and advocate for using Squared Error for Coverage (SEC) and Absolute Error for Coverage (AEC) metrics and empirical histograms of the coverage results as supplementary metrics. These enhancements provide a more comprehensive understanding of model performance in heterogeneous data-generating processes (DGPs). The practical implications are demonstrated through two examples, showcasing the benefits of these methodological improvements, which can significantly improve the robustness and accuracy of future research in statistical models for CATE and ATE estimation.

Additional Authors: Francisco Louzada

Abstract number: 104

Date: Thursday 10th April

Title: Shrinkage Bayesian Causal Forest with Instrumental Variable

Presenter: Jens Klenke

Affiliation: University of Duisburg-Essen

Topic: Heterogeneous treatment effects

Abstract: This paper proposes a novel framework for estimating heterogeneous treatment effects using Instrumental Variables (IV) in observational studies with sparse dataand imperfect compliance. Traditional IV methods, such as two-stage least squares(2SLS), often impose linearity assumptions that may not hold in complex empirical settings. To address these limitations, we build upon the Bayesian Instrumental Variable Causal Forest (BCF-IV) framework that has been developed to estimatethe conditional Complier Average Causal Effect (CACE) non-parametrically while retaining interpretability. BCF-IV uses Bayesian Additive Regression Trees (BART) to identify treatment effect heterogeneity and to estimate the conditional CACE based onthe conditional Intention-To-Treat (ITT) effects and the proportion of compliers. Our approach extends BCF-IV by proposing a Shrinkage Bayesian Instrumental Variable Causal Forest (SBCF-IV) algorithm. SBCF-IV adopts the SoftBART algorithm and makes two major contributions related to its shrinkage of splitting probabilities beyond shrinkage applied to individual tree contributions in BART. First, SBCF-IV implicitlydiscriminates between relevant and irrelevant covariates when estimating conditional ITT effects and proportions of compliers. Second, our approach implements varying posterior splitting probabilities into the discovery of heterogeneous subgroups. Thesemodifications improve SBCF-IV's ability to handle sparse data and to detect variables that drive the heterogeneity of treatment effects. A simulation study suggests a more precise estimation of conditional CACE while maintaining interpretability, particularlyin scenarios with sparsity, confounding, and nonlinearity.

Additional Authors: Lennard Maßmann

Abstract number: 105

Date: Thursday 10th April

Title: Targeting relative risk heterogeneity with causal forests

Presenter: Vik Shirvaikar

Affiliation: University of Oxford

Topic: Heterogeneous treatment effects

Abstract: State-of-the-art methods for estimating heterogeneous treatment effects, including causal forests (Wager and Athey, 2018), generally rely on recursive partitioning for non-parametric identification of relevant covariates and interactions. However, like many other methods in this area, causal forests partition subgroups based on differences in absolute risk. This can dilute statistical power by masking variability in the relative risk, which is often a more appropriate quantity of clinical interest. In this work, we propose and implement a methodology for modifying causal forests to target relative risk, using a novel node-splitting procedure based on exhaustive generalized linear model comparison. We present simulation results that suggest relative risk causal forests can capture otherwise undetected sources of heterogeneity. We additionally demonstrate the approach on real-world data from a large cardiovascular outcome trial, where subgroup effects in the primary endpoint and in secondary safety endpoints are both of clinical interest.

Additional Authors: Andrea Storas, Xi Lin, Chris Holmes

Abstract number: 106

Date: Thursday 10th April

Title: The Parachuted Hybrid CATE Estimator with Bootstrap Methods for Inference

Presenter: Xianlin Sun

Affiliation: The University of Hong Kong

Topic: Heterogeneous treatment effects

Abstract: In the current study, we introduce an innovative estimator for the Conditional Average Treatment Effect (CATE), hereafter referred to as the CATE estimator. This estimator is distinguished by its double robustness property, a characteristic ensuring consistency of the estimator provided that either the model for the propensity score or the model for the conditioned expected outcome is accurately specified. Notably, our estimator retains its consistency under the failure of both conditions, albeit exhibiting a reduced rate of convergence akin to that observed with non-parametric estimators. This attribute has led us to term our proposed estimator as the "parachuted" estimator, signifying its enhanced reliability and consistency even in scenarios where traditional assumptions regarding propensity scores or conditioned outcomes are not met. Furthermore, when compared to non-parametric estimators, our model demonstrates superior efficiency, offering a potentially higher rate of convergence under parametric conditions.To formulate this estimator, we employed a methodology that amalgamates a parametric approach based on Augmented Direct Learning (Meng and Qiao, 2022) with a non-parametric strategy utilizing kernel estimation (Abrevaya, Hsu, and Lieli, 2015). The integration technique, as proposed by Lee and Soleymani (2015), is instrumental in ensuring that our estimator not only achieves optimal convergence rates but also maintains consistency across evaluations.A pivotal achievement of this research is the derivation of the asymptotic distribution for this hybrid estimator. This distribution is characterized by having its mean aligned with the estimated target—the true CATE—and its variance describable through a closed-form expression. Moreover, our investigation extends into the statistical inference concerning our parachuted estimator for CATE via established bootstrap techniques. Through rigorous theoretical analysis grounded in the work of Chatterjee and Bose (2005), we provide substantive proof affirming that these bootstrap methods yield consistent estimations within our specified framework, subject to certain regular constraints.This advancement not only underscores the feasibility of integrating bootstrap methodologies into causal inference analyses specifically tailored for estimating CATE but also establishes a foundational theoretical basis for such applications. Consequently, our findings contribute significantly to the statistical literature by offering a novel approach for estimating CATE with enhanced robustness and efficiency, thereby extending the methodological toolkit available for causal inference research.

Additional Authors: Prof. Stephen M.S. Lee

Abstract number: 107

Date: Thursday 10th April

Title: Negative Control Outcomes for Selection Bias in Instrumental Variable Analysis

Presenter: Apostolos Gkatzionis

Affiliation: MRC Integrative Epidemiology Unit, University of Bristol

Topic: Instrumental variables/unmeasured confounding

Abstract: Instrumental variable (IV) analysis is a popular approach for causal inference. The instrumental variable design allows researchers to guard against confounding bias and reverse causation. However, selection bias remains an important concern for IV analyses. Negative controls are commonly used as a sensitivity analysis to detect biases in observational studies, but their use in IV analyses has been limited. In this talk, we explore the use of negative control outcomes to detect selection bias in IV analyses. Using causal diagrams, we briefly describe how selection bias affects IV studies. We then discuss under what conditions a variable can be used as a negative control outcome to detect selection bias in such studies. As with other sources of bias, we show that the main requirement is that the negative control outcome shares confounders with the original exposure and outcome. The effect of the negative control on study participation is of secondary concern; for example, a variable that does not affect participation can be a valid negative control for an IV outcome that does. We also argue that age and sex, sometimes used as negative control outcomes in IV analyses, are not valid negative controls in general but can still be useful under certain assumptions. In a real-data analysis, we investigate the pairwise causal relationships between 19 traits, utilizing data from the UK Biobank and using genetic variants as instruments. Treating biological sex as a negative control outcome, we identify selection bias in analyses involving commonly used traits such as alcohol consumption, body mass index and educational attainment.

Additional Authors: Kate Tilling

Abstract number: 108

Date: Thursday 10th April

Title: Estimating Heterogeneous Causal Effects with Tree-Based Methods under Imperfect Compliance and Positivity Violations

Presenter: Karolina Gliszczynska

Affiliation: University Duisburg-Essen

Topic: Instrumental variables/unmeasured confounding

Abstract: Estimating the Complier Average Causal Effect (CACE) in instrumental variable (IV) settings is critical for understanding causal relationships, particularly when treatment compliance varies across subpopulations. In real-world scenarios, implementing policies or interventions often results in imperfect compliance. Furthermore, real-world data usually contain regions where overlap is violated, posing significant challenges to standard estimation techniques and complicating policy or intervention effectiveness evaluation.This work extends the Bayesian Additive Regression Trees with Instrumental Variables (BART-IV) framework to a (Transformed) Random Forest-IV setting for estimating CACE. We introduce a methodology that incorporates kernel-based weighting to balance observable covariate distributions between groups defined by instrument assignment. This approach effectively addresses overlap violations and mitigates issues arising from extreme or near-deterministic instrument assignment probabilities.Our approach centers on two key strategies:Developing a Random Forest-IV framework that performs comparably to BART-IV while significantly reducing computational time—or even outperforming BART-IV when the binary covariate assumption is relaxed.Integrating kernel-based weights within the transformed Random Forest-IV framework to enhance the estimation of CACE.We demonstrate that both strategies perform well under varying levels of treatment effect heterogeneity. Notably, the kernel-weighted approach is particularly interesting in scenarios where propensity scores are close to 0 or 1, conditions that often lead to extreme inverse probability weights.

Additional Authors: .

Abstract number: 109

Date: Thursday 10th April

Title: Assessing changes in cell composition in observational single-cell gene expression studies

Presenter: Amber Huybrechts

Affiliation: Ghent University

Topic: Instrumental variables/unmeasured confounding

Abstract: Analysis of single-cell sequencing data, in particular cell abundance data where one counts the number of cells detected for each cell type, involves issues regarding data compositionality. Indeed, cell composition data contains only relative information on a cell type's abundance. An increase in one cell type might therefore also be reflected as a decrease in other cell types' abundance. This makes estimating causal disease effects in cell composition data rather complicated, especially in the presence of confounders. Using a case study, presented by Perez et al. (2022) [1], involving cell type abundance data of lupus patients from European and Asian ancestry, different methodologies are evaluated. Among these methods are LinDA and voomCLR, both of which were developed for investigating differential cel type abundance. These methods were compared to causal inference approaches that are designed to address confounders in a counterfactual framework; inverse probability weighting and standardization. Although both LinDA and voomCLR showed promising results regarding FDR control, voomCLR seems to be somewhat too conservative, resulting in lower sensitivity. We conclude that none of the aforementioned methods is sufficient to deal with both confounders and compositionality simultaneously. [1]Perez et al. Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. Science. 2022 Apr 8;376(6589):eabf1970. doi: 10.1126/science.abf1970. Epub 2022 Apr 8. Erratum in: Science. 2024 Jul 12;385(6705):eadr4064. doi: 10.1126/science.adr4064. PMID: 35389781; PMCID: PMC9297655.

Additional Authors: Koen Van den Berge, Sanne Roels, Oliver Dukes

Abstract number: 110

Date: Thursday 10th April

Title: Understanding spatial models with causal inference:

Presenter: Tim Lucas

Affiliation: University of Leicester

Topic: Instrumental variables/unmeasured confounding

Abstract: A wide array of statistical approaches have been developed for working with spatial data in a number of fields, including epidemiology, ecology and environmental science. Here we show how viewing these methods in a causal inference framework can make the aims, mechanisms and assumptions behind these methods clearer, making them easier to understand. Furthermore, this can make it easier to see connections and similarities between methods across disciplines. Even when causal inference is not the ultimate goal, many modelling applications require separating observation processes from the true mechanism of interest. This separation itself is inherently a causal problem and, as we show, describing existing approaches as such can make their rationale and limitations clearer. We intend this work to help guide practitioners in choosing a suitable method for their problem and applying it to estimate the desired quantities. Therefore we present the content in a non-mathematical format with a focus on applied research.

Additional Authors: Rohan Arambepola

Abstract number: 111

Date: Thursday 10th April

Title: E-value for RMTL: Evaluating Outcomes Under Non-Proportional Hazards

Presenter: Carmen Reep

Affiliation: Erasmus MC

Topic: Instrumental variables/unmeasured confounding

Abstract: IntroductionIn studies where cumulative incidence curves converge at follow-up, endpoint risk differences may overlook meaningful treatment effects. For example, in the WEAN SAFE switch study, earlier switches from controlled to assisted ventilation in mechanically ventilated patients were associated with earlier and/or more frequent successful extubations, reducing serious complications and ICU demand. In such scenarios, measures like restricted mean time lost (RMTL), which captures the area under the cumulative incidence curve, are more informative, with higher RMTL values indicating earlier or more frequent events (PMID: 34550319).To estimate the potential influence of unmeasured confounding in observational studies, the E-value offers a robust metric (PMID: 28693043). However, current E-value methods assume proportional hazards (PH), creating a gap for cases where PH does not hold, and risks converge over time. Here, we propose a novel approach to report and interpret E-values under these conditions.MethodsOur approach involves calculating E-values at each time point in the follow-up period and plotting them beneath the cumulative incidence curve. These E-values are then compared to the outcome risk factor of the strongest measured confounder. We report the mean E-value and the percentage of E-values exceeding the maximum observed confounder risk. This method is illustrated using data from the WEAN SAFE switch study.ResultsThe mean E-value over the 28-day follow-up was 2.3, with 71% of E-values exceeding the risk associated with the strongest measured confounder (P/F ratio, hazard ratio [HR] = 1.42). Clinical experts anticipate an unmeasured confounder stronger than the P/F ratio to be highly unlikely.ConclusionsFor observational studies mimicking randomized controlled trials, where PH assumptions are not met, RMTL provides a valuable measure when treatment effects manifest earlier in time. Evaluating the E-value across the follow-up period offers an innovative method to assess the potential influence of unmeasured confounding. Comparing these E-values to the strongest measured confounder and leveraging expert input enhances confidence in the findings.

Additional Authors: Evert-Jan Wils

Abstract number: 112

Date: Thursday 10th April

Title: Confounding of the competing event in time-to-event analyses

Presenter: Jost Viebrock

Affiliation: Department of Biometry and Data Management, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

Topic: Instrumental variables/unmeasured confounding

Abstract: Unmeasured confounding presents a fundamental challenge in observational studies. In many biomedical applications, the focus is on the time until an event of interest (e.g., death from breast cancer) occurs, which may be prevented by a competing event (e.g., death from other causes) occurring first. While unmeasured confounding of treatment and outcome of interest has been much explored, the specific impact of an unmeasured common cause affecting the treatment and the competing event—here referred to as competing confounding—has received little attention. For the total effect, a causal estimate in competing event settings introduced by Young et al. (2020), we examine how such competing confounding induces bias.

Additional Authors: Bianca Kollhorst, Vanessa Didelez

Abstract number: 113

Date: Thursday 10th April

Title: A promotion time cure model with endogeneity simultaneously affecting the cure probability and survival time of the susceptibles

Presenter: Lasse Winding

Affiliation: ORSTAT, KU Leuven

Topic: Instrumental variables/unmeasured confounding

Abstract: Competing confounding is particularly relevant when the competing event occurs much more frequently than the event of interest or when this confounding is strong. For example, consider assessing the effect of mammography screening on breast cancer mortality in the absence of information on smoking behavior. In this context, smoking acts as a potential unmeasured competing confounder, exerting a great influence on mortality from other causes (e.g., lung cancer) while being associated with participation in the mammography screening programme.

Additional Authors: Jean-Pierre Florens, Ingrid Van Keilegom

Abstract number: 114

Date: Thursday 10th April

Title: The role of estimation in Mendelian randomization: should Mendelian randomization investigations provide estimates?

Presenter: Benjamin Woolf

Affiliation: University of Bristol and University of Cambridge

Topic: Instrumental variables/unmeasured confounding

Abstract: We conduct a simulation study in a parametric setting (Weibull distributed event times) to illustrate how and when competing confounding leads to bias. We investigate the extent of bias under different scenarios, such as varying prevalence of the unmeasured confounder or its strength. Furthermore, we illustrate the potential structure of bias when the competing confounder additionally acts as a confounder for the event of interest. Moreover, we provide analytic formulas to calculate the bias under various parametric settings. These can be used, for instance, to investigate worst-case scenarios under different assumptions about the unobserved confounder. Moreover, they allow us to characterise when the total effect might be systematically over- or underestimated. For the mammography screening example, we provide an assessment of the potential bias under a variety of realistic settings.

Additional Authors: Benjamin Woolf, Stephen Burgess

Abstract number: 115

Date: Thursday 10th April

Title: Interventional effects modeling in telework-health research: exploring the indirect and direct effects in the telework work-engagement relationship

Presenter: Eduardo A. Bracho Montes de Oca

Affiliation: Sciensano, UCLouvain, UGent

Topic: Mediation analysis

Abstract: Background: Teleworking has increased exponentially over the past 50 years, yet its impact on work engagement – a key factor for workplace well-being - remains unclear. Work engagement is a key factor in workplace well-being. This study aimed to examine whether telework frequency is related to work engagement by doing an interventional mediation analysis, without assuming an underlying structure between mediators.Methods: Data from the working participants aged 18 to 65 years (n = 2,323) included in the June 2023 BelHealth survey (n = 7,186) were analyzed Telework frequency was categorized as: non-teleworkers (reference group), monthly, weekly, and full-time teleworkers (A). Hypothesized mediators were: emotional load, role conflict, workload, social support, autonomy, and skills-use (M, SIMPH questionnaire). Work engagement (Y, UWES-3) was the outcome. Two linear regression models were constructed per outcome (E(Y|A,M,C), E(M|A,C)), including M-M, M-M-A, and M-A interactions; while adjusting for confounding factors (C, e.g. education, age). Interventional direct and indirect effects were derived. Sensitivity analysis with 1,000 permutations assessed robustness, with 95% confidence intervals bootstrapped.Results: Weekly telework was associated in a decrease of work engagement (total effect) of 11.6% (-0.187; -0.043). Monthly (4.4%; -0.142; 0.052), and full-time telework (3.4%; -0.209; 0.132) showed a decrease as compared with non-teleworkers. Similarly, monthly telework (4.4%, -0.142; 0.052) was associated with a work engagement reduction. Full-time telework was associated to increase work engagement (5.3%, -0.498; 0.525). Direct effects of weekly telework reduced work engagement by 17.9% (-0.258; -0.097), while indirect effects showed reductions through emotional load (2.7%, -0.05; -0.006). Conversely, role conflict (+1.2%, 0.003; 0.027), social support (+2.2%, 0.008; 0.041), and skills-use (+5.7%, 0.019; 0.096) had indirect protective effects. Conclusion: This study highlights three key findings: 1) the frequency of telework differently impacts well-being at work based on the (in)direct effects, 2) telework frequency has a direct effect on work engagement, and 3) most mediators contributed indirectly to promoting work engagement.

Additional Authors: Robby De Pauw; Beatrijs Moerkerke; Barbara Cagnie; Lydia Gisle ; Bas de Geus

Abstract number: 116

Date: Thursday 10th April

Title: Accessible methods for causal mediation analysis in two-level data structures

Presenter: Kirsty James

Affiliation: King's College London

Topic: Mediation analysis

Abstract: Methods for causal mediation analysis are well established for non-hierarchical structures, but less so for two-level data. In this work we have reviewed accessible methods for mediation analysis in two-level structures with the intention of application to clinical trials data. In a two-level setting there are further quantities that can be estimated, and it might be possible to relax assumptions to estimate them. We will extend Rubin's causal modelling framework to define these causal estimands and assumptions. This review summarises methods using this consistent terminology. Multilevel mediation analysis is possible to implement in general purpose software. A common feature of many of the approaches is that they consider the hierarchical structure a "nuisance" leading to incorrect standard errors of the estimates if not accounted for, which is consistent with thinking about clustering on total treatment evaluation in trials (Krull and MacKinnon 2001). For continuous mediator and outcome variables a structural equation modelling framework has been put forward (Preacher, Zyphur et al. 2010). Assuming no interference such modelling defines separate direct and indirect effects at each level, a natural indirect effect between clusters (NIE(B)) and another natural indirect effect within clusters (NIE(W)).Other methods capitalise on the idea that within cluster correlation provides extra analysis information to relax assumptions (Talloen, Moerkerke et al. 2016). The natural indirect effect (NIE) can be further partitioned into a natural within indirect effect (NWIE) and a natural between indirect effect (NBIE). For continuous mediator and outcome, the NWIE can be estimated without bias even with unobserved mediator-outcome confounding at the cluster level. Individual participant data (IPD) meta-mediation analysis (Riley, Stewart et al. 2021) is a special case of multilevel mediation where the interventions, mediator and outcome vary at the individual level and the study defines the higher level. Approaches are categorised as "one-stage" or "two-stage". The more flexible two-stage approach calculates study-specific indirect effect estimates and then estimates the "expected NIE" by applying aggregate level meta-analysis methods. Promising methods are those that allow for interference and relaxation of the no mediator-outcome unobserved confounding assumption which is often difficult to satisfy in a clinical trial setting.

Additional Authors: Sabine Landau, Linda Sharples, Trudie Chalder, Richard Emsley

Abstract number: 117

Date: Thursday 10th April

Title: An adaptive test for natural indirect effect in large-dimensional mediation analysis

Presenter: Feng Liang

Affiliation: Gent University

Topic: Mediation analysis

Abstract: Detecting the absence of the mediation effect is a major focus of mediation analysis. When treatment does not influence mediators that do not affect the outcome, testing the natural indirect effect is an interesting issue, since the asymptotic distributions of existing test statistics vary under different sub-null hypotheses. This paper introduces a novel statistical inference procedure tailored for high-dimensional mediation structures to address the issue of test conservativeness under some nontrivial sub-null hypotheses. We first suggest a procedure using a partial penalized least squares estimation and compute the inner product of the treatment-mediator and the mediator-outcome coefficients. Based on the product, we develop a Wald-type test to handle the case where the mediator affects the outcome. When the mediators do not affect the outcome, the Wald-type test statistic fails to maintain the significance level. We then construct another test for the significance level maintenance. The final test is an adaptive-to-sub-null hybrid of the two tests, which can flexibly accommodate different sub-null hypotheses and ensures that the limiting null distributions converge to a common Chi-Square distribution uniformly across all sub-null hypotheses. Numerical studies are conducted to assess the finite sample performances of the proposed test and make comparisons with existing methodologies.

Additional Authors: Lixing Zhu

Abstract number: 118

Date: Thursday 10th April

Title: Longitudinal Mediation Analysis of Multiple Mediators: Estimating Brain Region-Specific Effects in Major Depressive Disorder

Presenter: Can Long

Affiliation: Section of Biostatistics, Department of Public Health, University of Copenhagen

Topic: Mediation analysis

Abstract: Current antidepressants for major depressive disorder (MDD) often show low remission rates and significant side effects, highlighting the importance of understanding the mechanisms behind treatment efficacy. Advances in neuroimaging, particularly BOLD fMRI, provide rich high-dimensional data on brain activity and connectivity, offering new opportunities to study how neural processes mediate treatment outcomes. However, it also poses challenges such as high dimensionality of mediators, temporal changes, and complex brain interactions, underscoring the need for more specialized mediation frameworks designed for imaging data.This work presents a longitudinal mediation analysis framework tailored for medical imaging data. First, it defines region-level mediators across multiple time points from high-dimensional brain imaging data, which form numerous dynamic interactive paths. It then estimates path-specific effects using linear structural equation modeling (LSEM) and aggregates relevant paths into region-specific effects using predefined selection matrices informed by anatomical and statistical criteria. Researchers can customize these selection matrices to examine specific groups of paths, such as how a brain region contributes over time or participates in specific processes. By enabling flexible path selection and aggregation, this method supports tailored analyses of region-specific effects under different hypotheses and criteria, allowing comparisons between aggregation approaches to better understand how brain regions influence outcomes over time.Simulation studies demonstrate that the framework accurately estimates indirect effects for brain regions under defined aggregation criteria. Unlike methods focusing on single-time-point region-specific effects, this approach, considering dynamic interactions critical for brain synergy modeling, emphasizes the estimation and interpretation of region-specific effects over time. By incorporating flexible path selection criteria, it provides researchers with a clear and interpretable way to test brain region-specific effects. This framework facilitates the exploration of MDD treatment mechanisms and could serve as a versatile tool for causal mediation analysis in high-dimensional data.

Additional Authors: Brice Ozenne, Theis Lange

Abstract number: 119

Date: Thursday 10th April

Title: Mediation analysis in longitudinal intervention studies with an ordinal treatment-dependent confounder

Presenter: Mikko Valtanen

Affiliation: University of Turku

Topic: Mediation analysis

Abstract: In interventional health studies, causal mediation analysis can be employed to investigate mechanisms through which the intervention affects the targeted health outcome. Identifying direct and indirect (i.e. mediated) effects from empirical data, however, becomes complicated if the mediator-outcome association is confounded by a variable itself affected by the treatment. Here, we investigate identification of mediational effects under such post-treatment confounding in a setting with a longitudinal mediator, time-to-event outcome and a trichotomous ordinal treatment-dependent confounder. We assume that the intervention always affects the treatment-dependent confounder only in one direction (monotonicity), leading to a special case of partial identification and show that the mediational effects are then identified up to a single (stratum-specific) sensitivity parameter. The monotonicity assumption can be assessed from empirical data, based on restrictions on the conditional distribution of the treatment-dependent confounder. We avoid pitfalls related to post-treatment conditioning by treating the mediator as a functional entity and defining the time-to-event outcome as a restricted disease-free time. In an empirical analysis, we use data from the Finnish Diabetes Prevention study containing a cohort of individuals at high risk of type 2 diabetes. We employ a parametric Bayesian approach to assess the extent to which the effect of a lifestyle intervention on avoiding type 2 diabetes is mediated through weight reduction, with other health-related changes acting as treatment-dependent confounders.

Additional Authors: Tommi Härkänen, Matti Uusitupa, Jaakko Tuomilehto, Jaana Lindström, Kari Auranen

Abstract number: 120

Date: Thursday 10th April

Title: Performance of G-computation estimators in randomized controlled trials with high-dimensional covariates

Presenter: Muluneh Alene Addis

Affiliation: Ghent University

Topic: Randomized trials

Abstract: In response to recent guidance from the U.S. Food and Drug Administration, adjustment for baseline co-variates has become popular in randomized controlled trials (RCTs) with the goal of improving efficiencyand power. However, when the sample size is small relative to the number of covariates, standard maximum likelihood estimators of treatment effects in generalized linear models (GLMs) may exhibit large bias. This study investigates whether G-computation estimators for marginal treatment effects in RCTs, which are known for their robustness to model misspecication in canonical GLMs, are also less sensitive to bias when covariates are high-dimensional.We review existing literature on G-computation under proportional asymptotic regimes, where the number ofcovariates grows with the sample size, providing a theoretical foundation for high-dimensional settings. Using Monte Carlo simulations, we compare the performance of various G-computation estimators based on linear and logistic outcome models. Specifically, we evaluate the impact of data-adaptive covariate selection and assess the effectiveness of debiasing methods tailored for high-dimensional covariates. In addition, we evaluate the impact of small sample correction techniques and sample splitting on the performance of standard errors. Our findings aim to guide the application of G-computation in modern RCTs, particularly when addressing challenges posed by high-dimensional data or small sample size.

Additional Authors: Kelly Van Lancker, and Stijn Vansteelandt

Abstract number: 121

Date: Thursday 10th April

Title: Estimating treatment effect in randomized trial after control to treatment crossover using external controls

Presenter: Christiana Drake

Affiliation: University of California, Davis

Topic: Randomized trials

Abstract: In clinical trials, it is common to design a study that permits the administration of an experimental treatment to participants in the placebo or standard of care group post primary endpoint. This is often seen in the open-label extension phase of a phase III, pivotal study of the new medicine, where the focus is on assessing long-term safety and efficacy. With the availability of external controls, proper estimation and inference of long-term treatmenteffect during the open-label extension phase in the absence of placebo-controlled patients are now feasible. Within the framework of causal inference, we propose several difference-in differences (DID) type methods and a synthetic control method (SCM) for the combination of randomized controlled trials and external controls. Our realistic simulation studies demonstrate the desirable performance of the proposed estimators in a variety of practical scenarios. In particular, DID methods outperform SCM and are the recommended methods of choice. An empirical application of the methods is demonstrated through a phase III clinical trial in a rare disease.

Additional Authors: Xiner Zhou

Abstract number: 122

Date: Thursday 10th April

Title: Powering RCTs for marginal effects with GLMs using prognostic score adjustment

Presenter: Emilie Højbjerre-Frandsen

Affiliation: Novo Nordisk A/S and Aalborg University

Topic: Randomized trials

Abstract: Estimating causal effects from randomized experiments is a fundamental aspect of clinical re-search, and enhancing the precision of these analyses is a key goal for statisticians. Historical data from registries, prior clinical trials, and health records offer a rich resource for understanding patient outcomes under standard-of-care and should be leveraged to increase study power. However, many existing methods for historical data utilization trade off reduced variance for less stringent type-I error rate control. This presentation presents a novel approach to leveraging historical data, specifically focusing on covariate adjustment for generalized linear models (GLMs) to enhance the efficiency of trial analyses without introducing bias. Our method involves training a prognostic model using historical data and then estimating the marginal effect using the plug-in GLM regression procedure proposed by Rosenblum & van der Laan in 2009, while adjusting for the trial subjects' predicted outcomes, known as their prognostic scores, within the linear predictor. This extends the approach of Schuler 2021 beyond the simple linear model. Under certain conditions, this prognostic score adjustment procedure achieves the minimum possible variance among a broad class of estimators. Even when these conditions are not fully met, prognostic covariate adjustment remains more efficient than raw covariate adjustment, with the efficiency gains depending on the prognostic model's predictive accuracy beyond the linear relationship with the raw covariates. We validate our approach through simulations and a reanalysis of a clinical trial conducted by Novo Nordisk A/S, demonstrating notable reductions in the variance of the marginal effect estimate. Finally, we present a simplified formula for asymptotic variance, facilitating power calculations that account for these efficiency gains.

Additional Authors: Alejandro Schuler, UC Berkeley

Abstract number: 123

Date: Thursday 10th April

Title: Federated Causal Inference: Multi-Site ATE Estimation beyond Meta-Analysis

Presenter: Rémi Khellaf

Affiliation: INRIA

Topic: Randomized trials

Abstract: In this work, we study Federated Causal Inference, an approach to estimate treatment effects from decentralized data across centers, or studies. We compare three classes of Average Treatment Effect (ATE) estimators derived from the Plug-in G-Formula, ranging from simple meta-analysis to one-shot and multi-shot federated learning, the latter leveraging the full data to learn the outcome model (albeit requiring more communication). Focusing on Randomized Controlled Trials (RCTs), we derive the asymptotic variance of these estimators for linear models. Our results provide practical guidance on selecting the appropriate estimator for various scenarios, including heterogeneity in sample sizes, covariate distributions, treatment assignment schemes, and center effects. We validate these findings with a simulation study.

Additional Authors: Aurélien Bellet, Julie Josse

Abstract number: 124

Date: Thursday 10th April

Title: Bayesian covariate-adjusted response-adaptive intervention allocation in stepped-wedge cluster randomised trials

Presenter: Constantin Schmidt

Affiliation: MRC Biostatistics Unit, University of Cambridge

Topic: Randomized trials

Abstract: Many real-world interventions are believed to be beneficial based on prior studies but face logistical constraints that prevent immediate implementation across all eligible units. In such cases, interventions are typically rolled out sequentially, with clusters of units crossing over into the intervention in a staggered fashion until all units are exposed. When the order of exposure is randomised before roll-out, the implementation design is referred to as a stepped-wedge cluster randomised trial (SW-CRT). The SW-CRT design facilitates rigorous evaluation of the intervention. Alternatively, the order of exposure might be chosen to maximise the benefit achieved by the intervention during the trial itself. We propose a novel trial design that balances the need for rigorous intervention evaluation with the objective of maximising benefit achieved by the intervention during the trial by incorporating Bayesian covariate-adjusted response-adaptive randomisation into the SW-CRT framework. During interim analyses, cluster- and patient-level covariates, along with accumulated response data, are used to predict intervention effects in yet-to-be-exposed clusters. Randomisation probabilities are then adjusted so that clusters with higher predicted intervention effects are more likely to receive the intervention earlier. Re-randomisation is carried out at each interim analysis in a way that respects the logistical constraints. Additionally, the trial can be initiated with unequal randomisation probabilities reflecting prior beliefs about intervention effectiveness. Future work will explore incorporating early stopping rules and implications for trial power.

Additional Authors: Pantelis Samartsidis (University of Cambridge); Lukas Pin (University of Cambridge); Shaun Seaman (University of Cambridge); Sofia Villar (University of Cambridge); and Daniela De Angelis (University of Cambridge)

Abstract number: 125

Date: Thursday 10th April

Title: Establishing when to use causal machine learning for conditional average treatment effect estimation in randomised controlled trials using simulation

Presenter: Eleanor Van Vogt

Affiliation: Imperial College London

Topic: Randomized trials

Abstract: Randomised controlled trials (RCTs) typically focus on estimating the average treatment effect (ATE), which often results in null conclusions. However, where heterogeneous treatment effects (HTEs) are of interest, factors responsible for variation must be pre-specified. There is growing interest in exploring HTEs in the context of personalised treatment regimens and policy decisions, and causal machine learning methods for HTEs are increasing in popularity. They offer flexible tools for exploring HTEs across many covariates without needing pre-specification by learning the conditional average treatment effect (CATE). Current usage of these methods is restricted to exploring HTEs and generating hypotheses for validation in a future dataset. Existing questions relate to the sample sizes required to obtain valid CATEs and the impact of missing covariate information on resulting CATEs. We conduct a simulation study to compare several causal machine learning candidates and classical subgroup detection methods across simple and complex HTE scenarios with varying sample sizes and missing data mechanisms. We consider binary and continuous outcomes and the handling of competing events. We explore bias, coverage of HTE estimates, and error rates for global heterogeneity tests.Informed by minimum sample size requirements from our simulation and results from heterogeneity testing, we additionally simulate scenarios where HTE hypotheses are generated during an interim analysis of an RCT and then validated on the later recruited participants. Large RCTs could potentially use this approach to generate and validate subgroup findings and provide treatment recommendations.By addressing challenges such as minimum sample size requirements and missing data handling, the presented results from simulations will provide researchers with a framework to decide whether causal machine learning methods are suitable for RCT datasets at their disposal. Further, the proposed interim analysis framework has the potential to enhance RCT utility, enabling real-time hypothesis generation and validation for personalised, evidence-driven treatment.

Additional Authors: Karla Diaz-Ordaz, Suzie Cro