EuroCIM 2023, Oslo

Abstracts

Invited presentation 1

Target trial emulation for policy evaluation

***Bianca De Stavola**, University College London; Vincent Nguyen, Kate Lewis, Lucy Karwatowska, Lorraine Dearden*

The talk addresses the challenges of designing studies that address policy evaluation questions using linked administrative health and education data. This talk is motivated by an investigation of the health impact of special educational needs (SEN) provision on children with medical and healthcare needs in the UK from 2006 to 2019. Designing such a study highlights the difficulties of defining the target population (children in need of SEN provision, a very heterogeneous group), qualifying the exposure (school-recorded SEN), understanding its drivers whilst also accounting for clustering of provision by local authority and school. Additionally, data completeness and temporal variability in coding practices raise issues of how to account for selection bias (due to selection into the study) and residual confounding (due to poorly measured confounders). Adopting a target trial emulation framework in this context helps clarify the limitation of the data and the questions that can be addressed. The talk will discuss possible estimands of interest. Results obtained by alternative estimation approaches will be compared, exploiting methods from the biostatistics and econometrics literature and the richness of the individual level data, either controlling for residual confounding (e.g., using g-computation) or exploiting the observed local authorities' variation in timing and extent of provisions (using IV or difference-in-difference).

Invited presentation 2

Bespoke instrumental variables for causal inference

***Oliver Dukes**, Ghent University; David Richardson, UC Irvine; Zach Shahn, CUNY; Eric Tchetgen Tchetgen, UPenn*

Many proposals for the identification of causal effects in the presence of unmeasured confounding require an instrumental variable or negative control that satisfies strong, untestable exclusion restrictions. In this talk, I will instead show how one can identify causal effects for a point exposure by using a measured confounder as a 'bespoke instrumental variable'. This strategy requires an external reference population that does not have access to the exposure, and a stability condition on the confounder-outcome association between reference and target populations.

Invited presentation 3

Time-dependent mediators in survival analysis

Odd O. Aalen, *University of Oslo*

We discuss causal mediation analyses for survival data. The emphasis is on a dynamic point of view, that is, understanding how the direct and indirect effects develop over time. Mediation analysis in a survival setting is a challenge. One issue is that patients may survive in one counterfactual setting, and not in the other, or they may be censored in one and not in the other. An additional aspect is the need for understanding the development over time. The relevant mediators will typically be stochastic processes.

An approach to mediation analysis is based on nested counterfactuals. However, this implies a cross-world assumption which has been disputed by several authors. A simpler and more intuitive procedure was recently developed by Vanessa Didelez, based on earlier work by Robins and Richardson. The idea is to consider different components of the treatment; one component describes how the treatment affects the outcome through the mediator, while another describes the other effects of treatment.

To illustrate and evaluate assumptions we use causal graphs for local independence. This is a concept for describing how the evolution of a stochastic process depends on other processes. In comparison with conditional independence, local independence has the advantage that there is a direction in the dependence relationship. In this talk, we shall focus on a time-discrete version of local independence. We shall show how local independencies can be read off from the graph using δ-separation (which is an extension to stochastic processes of d-separation).

Invited presentation 4

Causal inference with random assignment versus researcher created binary treatments

Kenneth A. Bollen*, University of North Carolina*

The causal inference methods of Potential Outcomes (POs), Directed Acyclic Graphs (DAGs), and Structural Equation Models (SEMs) have contributed much to our understanding of causal effects. Yet the teaching and application of these methods (especially POs and DAGs) have nearly always regarded treatment as binary even when the magnitude of treatment can differ greatly. The two most common types of binary treatments are those from randomized experiments and those that are categorized versions of continuous treatments. Binary treatments via categorization are far more common in observational studies. I derive results showing that binary treatment variables that have different origins should be treated differently. Not doing so makes biased causal inferences more likely. I illustrate the value of combining POs, DAGs, and SEMs perspectives to illuminate potential problems with binary treatments rather than relying only on one perspective. The new analytic results are illustrated with simulations and an empirical example. Finally, I make recommendations on how researchers should analyze binary treatments.

Invited presentation 5

An alternative "interventionist causal inference": Leading with stories

Jessica Young, *Harvard Medical School*

The modern causal inference literature might suggest that researchers must approach causal inference from one of two perspectives: "interventionist" as characterized by a "target trial emulation framework" versus "noninterventionist" as characterized by Pearl's NPSEM-IE. A widely discussed distinction between these perspectives is the classification of the consistency condition as an assumption that may be violated when counterfactual outcomes are ill-defined ("interventionist") versus a definition guaranteed by the causal model ("noninterventionist"). However, implicit within the popular framing of both perspectives seems to be an assumption that the treatment of underlying interest to the investigator is measured in the data set to be analyzed. In this talk, we consider a third "perspective" that begins to more clearly emerge when we acknowledge the reality that most researchers are not trained in the "art" of articulating causal questions separate from the data in hand. A framework for "interventionist" causal inference capturing a broader set of counterfactual targets than the current understanding of a "target trial framework" is outlined. This alternative framework, which captures any counterfactual target contained within Robins's FFRCISTG model, is inspired by Robins's and Richardson's insights on treatment decompositions illuminated by investigator-led "stories" and aspects of Petersen's and van der Laan's notion of a "causal road map", emphasizing clear separation of questions from data and statistics. We pose these ideas in the context of the widely-debated example of an investigator stating interest in estimating an "effect of obesity" in observational data as well as causal effect estimation in actual randomized trials with truncation by death.

Invited presentation 6

Risk prediction under hypothetical interventions

***Ruth Keogh**, London School of Hygiene & Tropical Medicine; Daniala Weir, Utrecht University; Nan van Geloven, Leiden University Medical Center*

Clinical risk prediction models enable predictions of a person's risk of an outcome given their observed characteristics. It is often of interest to use risk predictions to inform whether a person should initiate a particular treatment. However, when standard clinical prediction models are developed in a population in which patients follow a mix of treatment strategies, they are unsuitable for informing treatment decisions. Risk prediction under hypothetical interventions aims to address this problem by providing estimates of what a person's risk would be if they were to follow a particular treatment strategy. This talk will discuss methods for development and validation of models for prediction under interventions, with a focus on use of longitudinal observational data to predict risk under interventions that are sustained treatment strategies. Methods from the literature that can be used for the task of prediction under interventions will be discussed, with a longer example focusing on use of marginal structural models. An essential step in development and reporting of any prediction model is to validate its performance. However, methods for validating standard clinical risk prediction models do not apply to interventional prediction models. I will present newly developed methods for assessing the predictive performance of interventional prediction models, including measures of discrimination and calibration. The methods will be illustrated using a new open-access synthetic data designed for assessment and comparison of causal inference methods. The data is designed to mimic real world data and is built around a case-study about treatments for type-2 diabetes.

Contributed presentation 1

Inference procedures in target trial emulation with survival outcomes: comparison of confidence intervals based on the sandwich variance estimator and two forms of bootstrap

Juliette Limozin, *MRC Biostatistics Unit, University of Cambridge; Li Su, MRC Biostatistics Unit, University of Cambridge; Shaun Seaman, MRC Biostatistics Unit, University of Cambridge*

The target trial emulation (TTE) framework proposed by Hernán et al. (2008, 2016) is a popular approach for providing causal inference of treatment effects from observational data. In TTE, inverse probability weighting is often used to adjust for time-varying confounding and/or dependent censoring (Hernán & Robins, 2000). Then structural models for potential outcomes can be fitted to the weighted data to obtain per-protocol effect estimates in emulated trials. For inference, the sandwich variance estimator with fixed weights is a popular but often conservative estimator for the variance of the structural model parameters. Motivated by nonparametric bootstrap's ability to account for variation in weight estimation, as well as linearised estimating function (LEF) bootstrap's computational efficiency (Hu & Kalbfleisch 2000, Binder & Roberts 2009), we compared the performance of confidence intervals (CIs) based on nonparametric bootstrap, LEF bootstrap and the sandwich estimator for the marginal risk difference of a survival outcome in various settings of sequential emulated trials (Danaei et al., 2013) through simulations. Our initial results indicate that with small/moderate sample sizes and not too imbalanced treatment groups, nonparametric bootstrap-based CIs had better coverage whereas sandwich estimator-based CIs' coverage deteriorated at later follow-ups. With large sample sizes, LEF bootstrap-based CIs performed similarly to nonparametric bootstrap-based CIs, and slightly better when treatment imbalances were large, while they were computationally much faster to construct. We are currently investigating the impact of the outcome event rate on these CI methods' performances.

Contributed presentation 2

Estimand strategies for safety outcomes

Alessandra Mattei, University of Florence; ***Fabrizia Mealli****, University of Florence*

Safety evaluation of new therapies is an essential aspect of clinical trials, with the primary focus of quantifying the incidence of adverse events (AEs) and comparing it to a standard treatment. Several estimand strategies have been proposed for efficacy analysis of time-to-event outcomes in the presence of censoring and competing events, also in accordance to the ICH E9 Addendum. Safety analysis of adverse events, instead, is often rather simplistic: AE probabilities are estimated without explicitly defining the target causal comparison and neglecting assumptions on the censoring mechanisms leading to differential follow-up times (e.g., in oncology trials, patients may discontinue the control treatment earlier than the new treatment due to Progressive Disease). Here, we explicitly define the assumptions under which estimators typically used in the literature, such as the Exposure-Adjusted Incidence Rate, Kaplan-Meier and Aalen-Johansen estimators, have a causal interpretation. We introduce new principal stratum and hypothetical estimand strategies for safety outcomes in the presence of censoring, competing events and varying follow-up times. We also propose identifying assumptions as well as estimators under these assumptions. Our contribution will enhance interpretation of AE risks and has the potential of changing clinical trial practice with regard to safety analysis and risk-benefit assessment.

Contributed presentation 3

Balanced and robust randomized treatment assignments: The finite selection model for the health insurance experiment and beyond

Ambarish Chattopadhyay, Stanford University; Carl Morris, Harvard University; ***Jose Zubizarreta****, Harvard University*

The Finite Selection Model (FSM) was developed by Carl Morris in the 1970s for the design of the RAND Health Insurance Experiment (HIE), one of the largest and most comprehensive social science experiments conducted in the U.S. In the FSM, a treatment group at each of its turns selects the available unit that maximally improves the combined quality of its resulting group of units according to a common optimality criterion. In the HIE and beyond, we revisit, formalize, and extend the FSM as a general tool for experimental design. Leveraging the idea of D-optimality, we propose and analyze a new selection criterion in the FSM. The FSM using the D-optimal selection function has no tuning parameters, is affine invariant, and can retrieve classical designs such as randomized block and matched-pair designs. For multi-arm experiments, we propose algorithms to generate a selection order of treatments. We demonstrate FSM's performance in a case study based on the HIE, a simulation study, and in ten randomized studies from the health and social sciences. On average, the FSM achieves 68% and 56% better covariate balance than complete randomization and rerandomization in a typical study. We recommend the FSM be considered in experimental design for its conceptual simplicity, efficiency, and robustness.

Contributed presentation 4

Shrinkage estimation for causal inference and experimental design

Evan Rosenman, *Harvard University; Luke Miratrix, Harvard University; Art Owen, Stanford University; Mike Baiocchi, Stanford University; Guillaume Basse, Citadel Securities*

How can increasingly available observational data be used to improve the design and analysis of randomized controlled trials? One approach is to couple an RCT with an observational study using shrinkage estimation. We operate in a stratified setting, and consider two key questions: 1) how can we develop shrinkage estimators that combine causal estimates from observational and experimental sources, and 2) using these estimators, how might we design experiments more efficiently?

To answer the former question, we extend results from Stein shrinkage. We propose a procedure for deriving shrinkage estimators that leverage observational and randomized data together, making use of a generalized unbiased risk estimate. We develop two new estimators, prove finite sample conditions under which they have lower risk than an estimator using only experimental data, and show that each achieves a notion of asymptotic optimality. We also draw connections between our approach and results from sensitivity analysis.

We next consider designing a prospective RCT. If we intend to shrink the experiment's causal estimates toward those of a completed observational study, how do we optimize the experimental design? We show that the risk of the shrinkage estimator can be computed via numerical integration, and propose algorithms for determining the best allocation of units to strata, accounting for the imperfect parameter estimates we would obtain from the observational study. Lastly, we show to impose "guardrails" on the design, so the experiment can be reasonably analyzed with and without shrinkage.

Contributed presentation 5

Causal inference with misspecified interference structure

***Daniel Nevo**, Tel Aviv University; Bar Weinstein, Tel Aviv University*

As an alternative to the no-interference assumption, an interference structure is often represented using a network. Ubiquitously, the network structure is assumed to be known and correctly specified. However, correctly encoding the interference structure in a network can be challenging. For example, edge might be missing, network structure can change over time, and contamination between clusters might be present. Using the exposure-function framework, we quantify the bias of commonly used estimators when the network interference structure is misspecified. Then, to overcome the problem of network misspecification, we propose two solutions. First, we propose a network misspecification-robust estimator utilizing multiple networks simultaneously which is unbiased if one of the networks correctly represents the interference structure. As an alternative, we propose a sensitivity analysis framework that quantifies the impact of a postulated interference structure misspecification on the causal estimate as a function of parameters governing a misspecification mechanism. We illustrate the bias arising from incorrectly specified network and study the bias-variance tradeoff entailed in our proposed misspecification-robust estimator. We demonstrate the utility of our methods in two real examples involving two different interference structures: a social network field experiment and a cluster-randomized trial.

Contributed presentation 6

Machine learning for difference-in-differences with staggered treatment timing and heterogeneous treatment effects

***Julia Hatamyar**, University of York, Centre for Health Economics; Noemi Kreif, University of York Centre for Health Economics; Martin Huber, University of Fribourg; Rudi Rocha*

This paper combines two rapidly developing areas of research in econometrics: machine learning (ML) for estimating heterogeneous treatment effects, and difference-in-differences (DiD) with staggered treatment adoption, ie. when units become treated at different times and remain treated thereafter. By embedding a non-parametric, ML-based method for estimating heterogeneous treatment effects in two-period DiD (Lu, Nie & Wager 2019) within the multiple time periods DiD framework of Callaway & Sant'Anna (2021), we are able to estimate time-varying conditional average treatment effects at the observation level. We also aggregate the individual treatment effects to study overall average treatment effects on the treated (ATT), and "dynamic" treatment effects - the ATT in each sequential post-treatment period. We make use of simulations to examine the performance of our proposed technique, and then apply it to the evaluation of the Family Health Programme (FHP) effect on infant mortality in Brazil. Both our simulations and Brazilian Family Health Programme application show similar performance of our proposed estimator to the group-time ATT of Callaway & Sant'Anna (2021). The dynamic ATTs are in line with previous work and show an average decline in infant mortality in the initial years after a municipality adopted FHP. Our estimates demonstrate a large degree of heterogeneity in this response across municipalities, especially over time - i.e., not all municipalities experienced the reduction in mortality. We explore how evolution of this treatment effect heterogeneity differs according to various measures of inequality, poverty and political affiliation across individual Brazilian municipalities.

Contributed presentation 7

Doubly robust machine learning for an instrumental variable study of surgical care for cholecystitis

***Luke Keele**, University of Pennsylvania; Kenta Takatsu, Alex Levis, Edward Kennedy, CMU*

Instrumental variables are widely used in econometrics and epidemiology for identifying and estimating causal effects when an exposure of interest is confounded by unmeasured factors. Standard IV estimation methods, however, are built on parametric models that are prone to bias from model misspecification. We develop a set of IV estimation methods based on the doubly robust machine learning framework. These estimation methods allow for a variety of machine learning methods to be used for nuisance parameter estimation, while being doubly robust with fast rates of convergence for honest inference. Specifically, we construct a regular asymptotic linear estimator under a nonparametric statistical model using semiparametric theory. We outline the use of sample splitting to ensure parametric convergence rates and honest inference. We use these methods for estimation of the primary target causal estimand in an IV design. We also extend these methods by developing estimators for a key sensitivity analysis, profiling principal strata, and heterogenous causal effects. We conduct a simulation study to demonstrate when more our more flexible estimation methods outperform status quo methods. We apply these methods to an application in comparative effectiveness research. Here, we study the treatments for emergency cholecystitis–inflammation of the gallbladder. One treatment for cholecystitis is removal via surgery. Alternative non-surgical treatments include managed care and treatment via pharmaceutical options. Randomization is judged to violate the principle of equipoise, so we use an instrument for operative care–the surgeons tendency to operate.

Contributed presentation 8

Instrumental variable approaches for addressing unmeasured confounding in estimating time-varying treatment effects

***Daniel Tompsett**, UCL; Stijn Vansteelandt, Ghent University; Richard Grieve, London School of Hygiene and Tropical Medicine; Manuel Gomes, UCL*

Time-varying confounding is a key challenge in comparative effectiveness of treatment strategies sustained over time, i.e. time-varying treatments. Methods for handling time varying-confounding are relatively well established but often make the 'no unobserved confounding' assumption. Instrumental variables (IV) can help address the unmeasured confounding but it's unclear how they perform in settings with time-varying confounding. This paper extends and critically evaluates two alternative approaches for incorporating time-varying IV in 1) inverse probability weighting-based marginal structural models (IV-MSM), and 2) structural nested mean models (IV-SNMM) via g-estimation. This study is motivated by an evaluation of sustained biologic treatment over 5 years of patients with severe rheumatoid arthritis, using data from the US National Databank for rheumatic diseases. We considered two potential time-varying IVs: medical prescribing preferences, and health insurance. We found that these IVs had different strengths of association with biologic treatment and variability over time. Irrespective of the IV used, IV-SNMM led to a different 5-year treatment effect on patient's health related quality of life and costs, and narrower confidence intervals compared to IV-MSM. Through a comprehensive simulation study we assessed the performance of these methods across realistic scenarios, including varying strengths of association between the IV and treatment over time, sample sizes and strength of the time-varying confounding. We find that IV-SNMM performs well throughout, even in scenarios with small sample sizes, strong confounding and 'weak' IV. IV-MSM only performs well in ideal circumstances, such as a 'strong' IV and large sample sizes and requires further development.

Contributed presentation 9

Instrumental processes using integrated covariances

Søren Wengel Mogensen*, Lund University*

Instrumental variable (IV) methods have a long history in several quantitative fields, including causal inference. IV methods leverage exogeneity of an instrumental variable to obtain consistent estimation of a confounded causal effect. In many applications in health science and economics, data consists of multivariate time series. In this case, standard IV methods do not apply due to correlation between variables at different time points. In time series data, one may use conditional IV methods to leverage exogeneity of an instrumental process for causal effect estimation in the presence of unobserved confounder processes. However, in this talk we take a different approach. We introduce a novel IV method in stationary time series by considering an integrated measure of covariance between the coordinate processes. This avoids the use of conditional IV methods and provides a conceptually simple solution to the time series IV problem in a quite broad class of time series. Moreover, this method can be extended to certain classes of continuous-time processes. Thus, it provides a unified IV approach in discrete-time and continuous-time multivariate stochastic processes. The proposed method is translation-invariant in the sense that estimation will be consistent even if each coordinate process is shifted in time by an unknown, process-specific quantity, creating a misalignment of the observed processes. We briefly discuss an application to a real alarm network where this is a very useful property as each alarm is known to be sampled with some lag and this lag is different and unknown for each type of alarm.

Contributed presentation 10

A cautious approach to constraint-based causal model selection based on equivalence tests

***Daniel Malinsky**, Department of Biostatistics, Columbia University*

Causal graphical models are used many scientific domains to represent important causal assumptions about the processes that underlie collected data. The focus of this work is on graphical structure learning (a.k.a. causal discovery or model selection) for the "downstream" purpose of using the estimated graph for a subsequent causal inference tasks, such as establishing the identifying formula for some causal effect of interest and then estimating it. An obstacle to having confidence in existing procedures in applied health science settings is that they have a tendency to estimate structures that are overly sparse, i.e., missing too many edges. However, statistical "caution" (or "conservativism") would err on the side of more dense graphs rather than more sparse graphs. This paper proposes to reformulate the conditional independence hypothesis tests of classical constraint-based algorithms as equivalence tests: test the null hypothesis of association greater than some (user-chosen, sample-size dependent) threshold, rather than test the null of no association. We argue this addresses several important statistical issues in applied causal model selection and leads to procedures with desirable behaviors and properties.

Contributed presentation 11

Improving causal discovery with temporal background knowledge

Christine Bang, *University of Bremen and Leibniz Institute for Prevention Research and Epidemiology – BIPS; Janine Witte, Leibniz Institute for Prevention Research and Epidemiology – BIPS; Ronja Foraita, Leibniz Institute for Prevention Research and Epidemiology – BIPS; Vanessa Didelez, University of Bremen and Leibniz Institute for Prevention Research and Epidemiology – BIPS*

Causal discovery aims to estimate a (causal) graph from data. These methods have well-known issues: The output in form of an estimated equivalence class (represented by a so-called CPDAG) can be sensitive to statistical errors and is often not very informative. As we address in this work, including background knowledge, if correct, can only improve (and never harm) the result of causal discovery. Here, we focus on temporal background knowledge as would be available in longitudinal or cohort studies. This type of knowledge is reliable, relatively straightforward to incorporate, and as we show show, the resulting estimated graphs have desirable properties. First, we propose an algorithm to fully incorporate temporal background knowledge, and we prove that this algorithm is sound and complete, and that the output is stable. We show that this algorithm outputs restricted equivalence classes (represented by so-called temporal MPDAGs) that are more informative, and more robust to statistical errors compared to CPDAGs. We illustrate the usefulness of this algorithm in practice by applying it to data from a cohort study investigating children's health from infancy to young adolescence. Second, we are able to characterise the temporal MPDAGs as distinct from those based on other types of background knowledge. Thus, we can determine precisely when temporal knowledge adds information, and when it is redundant. Finally, we show that this class of graphs inherits key properties of CPDAGs so that that the usual reasoning with graphs is retained, e.g. rules to find sufficient adjustment sets.

Contributed presentation 12

A longitudinal causal graph analysis investigating modifiable risk factors and obesity in a European cohort of children and adolescents

***Ronja Foraita**, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany; Janine Witte, Leibniz Institute for Prevention Research and Epidemiology ; Claudia Börnhorst, Leibniz Institute for Prevention Research and Epidemiology; Wencke Gwozdz, Department of Consumer Research, Communication and Food Sociology, Justus-Liebig-University; Valeria Pala, Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano; Lauren Lissner, School of Public Health and Community Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg; Fabio Lauria, Institute of Food Sciences, CNR, Avellino, Italy; Lucia Reisch, Department of Management, Society and Communication, Copenhagen Business School; Dénes Molnár, Department of Paediatrics, Medical School, University of Pécs; Stefaan De Henauw, Department of Public Health and Primary Care, Faculty of Medicine and Health Sciences, Ghent University; Luis Moreno, GENUD (Growth, Exercise, Nutrition and Development) Research Group, University of Zaragoza; Toomas Veidebaum, Department of Chronic Diseases, National Institute for Health Development, Tallinn, Estonia; Michael Tornaritis, Research and Education Institute for Child Health, Strovolos, Cyprus; Iris Pigeot, Leibniz Institute for Prevention Research and Epidemiology; Vanessa Didelez, Leibniz Institute for Prevention Research and Epidemiology*

In this work, we estimate a Cohort Causal Graph (CCG) over the life-course from childhood to adolescence, including lifestyle and health factors. We aim at identifying potential early causes of obesity to help determine promising targets for prevention strategies. We adapt the popular PC-algorithm to deal with missing values by multiple imputation, with mixed discrete and continuous variables, and with prior knowledge such as time-ordering. The new temporal PC-algorithm (tPC) outputs a maximally oriented partially directed acyclic graph (MPDAG), which is similar to a completed partially directed acyclic graph (CPDAG) but can contain more directed edges due to prior temporal knowledge. We apply the tPC to estimate causal relations among 51 variables including obesity, diet, early life factors, lifestyle, and cultural background of 5,112 children from the European IDEFICS/I.Family cohort across three waves (2007-2014). The resulting CCG suggests some but not many and only indirect possible pathways from earlier modifiable risk factors such as audio-visual media consumption (AVM) to later obesity. We will discuss the robustness of the graph based on further analyses such as using test-wise deletion and the bootstrap.

Contributed presentation 13

Causal discovery from tail, variance or other characteristics

***Juraj Bodik**, University de Lausanne; Valérie Chavez-Demoulin, University de Lausanne*

Estimating a causal structure from observational data is an important goal in many scientific fields. However, typical methods rely on strong assumptions about the data-generating process, such as the additivity of the noise in the structural equation models. These assumptions are the main drawback when using causal discovery methods in practice, since they are often uninterpretable and not testable. In this work, we introduce the Conditionally Parametric Causal Model (CPCM), which aims to provide a framework requiring a more realistic set of assumptions. The drawback is that it can be used only in certain situations when the effect arises from some limit theorem. In that case, we assume some shape of the conditional distribution, and we model only specific characteristics of the distribution, such as tail or variance. We provide theoretical results to establish the conditions under which the causal structure is identifiable using CPCM. We also propose an algorithm for estimating the causal structure from a random sample. As an example, we are interested in causal relations between certain variables describing salaries. In such a situation, it may be reasonable to assume that the effect given the cause is Pareto distributed with the parameter being a function of the cause, i.e. $Y|X \sim Pareto(\theta(X))$. As we show in our work, this assumption is strong enough to distinguish between the cause and the effect.

Contributed presentation 14

Policy learning with rare outcomes and other complications

Julia Hatamyar, University of York; ***Noemi Kreif****, University of York*

Machine learning (ML) estimates of conditional average treatment effects (CATE) can be used to inform policy allocation rules, by either treating those with a beneficial estimated CATE ("plug in policy"), or by searching for a decision tree that optimises overall outcomes. The advantage of the latter is its better interpretability and ability to discard undesirable (e.g. sensitive) variables from the targeting. Little is known about the practical performance of these algorithms in usual settings of policy evaluations. We contrast the performance of various policy learning algorithms, using synthetic data with varying outcome prevalence (rare vs. not rare), positivity violations, the extent of treatment effect heterogeneity and the sample size. From the ML models we form plug-in and tree-based policies, and compare the performance of the estimated policy allocation by its so-called regret: how far the benefit from a resulting policy allocation is from the best possible ("oracle" policy). We find that the plug-in policy type outperforms tree-based policies in terms of regret, regardless of ML method used. Within either policy class, Causal Forests and the Normalised-Double-Robust Learner performed best, while Bayesian Additive Regression Trees performed worst. Additionally, we find evidence that with small sample sizes or in settings where the ratio of covariates to samples is high, learning policy trees using CATEs has a better performance than using doubly-robust scores, while this difference disappears with sample size. The methods are applied to a case study that investigates infant mortality through improved targeting of subsidised health insurance in Indonesia.

Contributed presentation 15

From data to decisions: how effects of intervening variables can guide policies

Mats J. Stensrud, *Ecole Polytechnique Federale de Lausanne (EPFL); Lan Wen, University of Waterloo; Aaron Sarvet, EPFL*

This work is motivated by two major threats to valid causal inference: unmeasured confounding and ill-defined interventions. I will present new results on average causal effects in settings with unmeasured exposure-outcome confounding. These results are used to study a class of estimands that are currently not targeted by standard approaches, representing queries about the effect of a so-called intervening variable. As an example, I consider the role of chronic pain and doctors' prescriptions of opioids, and illustrate how conventional approaches will lead to estimates with ambiguous policy implications. Yet, the effects of intervening variables have clear policy implications and furthermore are non-parametrically identified by the classical frontdoor formula. As an independent contribution, I will introduce a new semiparametric efficient estimator of the frontdoor formula that comes with a uniform sample boundedness guarantee. The methodological results are used to estimate effects of doctors' opioid prescription policies on mortality.

Contributed presentation 16

Causal mediation analysis with multiple time-varying mediators

Sheng-Hsuan Lin, *Institute of Statistics, National Yang Ming Chiao Tung University*

In longitudinal studies with time-varying exposures and mediators, the mediational g-formula is an important method for the assessment of direct and indirect effects. However, current methodologies based on the mediational g-formula can deal with only one mediator. This limitation makes these methodologies inapplicable to many scenarios. Hence, we develop a novel methodology by extending the mediational g-formula to cover cases with multiple time-varying mediators. We formulate two variants of our approach that are each suited to a distinct set of assumptions and effect definitions and present nonparametric identification results of each variant. We further show how complex causal mechanisms (whose complexity derives from the presence of multiple time-varying mediators) can be untangled. A parametric method along with a user-friendly algorithm was implemented in R software. We illustrate our method by investigating the complex causal mechanism underlying the progression of chronic obstructive pulmonary disease. We found that the effects of lung function impairment mediated by dyspnea symptoms and mediated by physical activity accounted for 13.7% and 10.8% of the total effect, respectively. Our analyses thus illustrate the power of this approach, providing evidence for the mediating role of dyspnea and physical activity on the causal pathway from lung function impairment to health status.

Contributed presentation 17

On the causal interpretation of randomized interventional indirect effects

Caleb Miles, *Columbia University*

Identification of standard mediated effects such as the natural indirect effect relies on heavy causal assumptions. By circumventing such assumptions, so-called randomized interventional indirect effects have gained popularity in the mediation literature. Here, I introduce properties one might demand of an indirect effect measure in order for it to have a true mediational interpretation. For instance, the sharp null criterion requires an indirect effect measure to be null whenever no individual-level indirect effect exists. I show that without stronger assumptions, randomized interventional indirect effects do not satisfy such criteria. I additionally discuss alternative causal interpretations of such effects.

Contributed presentation 18

Characterizing immune correlates of protection in vaccine efficacy trials with stochastic-interventional causal effects

***Nima Hejazi**, Harvard T.H. Chan School of Public Health; Peter Gilbert, Fred Hutchinson Cancer Center; Mark van der Laan, University of California, Berkeley; David Benkeser, Emory University*

In clinical trials randomizing participants to active vs. control conditions and following study units until the occurrence of a primary clinical endpoint, evaluating the efficacy of a quantitative exposure or mediator is challenging. This is due, in part, to the fact that statistical innovations in causal inference have historically focused on defining interpretable estimands compatible only with categorical (or binary) treatments. We will introduce stochastic-interventional causal effects, which provide a measure of the effect attributable to perturbing a treatment's natural value, focusing primarily on how these effect definitions provide a scientifically informative solution when working with quantitative, continuous-valued intervention variables. Unfortunately, the estimation of these, and other, estimands in treatment or vaccine efficacy clinical trials often requires significant additional care, for such trials measure immunologic biomarkers – critical to understanding the mechanisms by which vaccines confer protection or as surrogate endpoints in future clinical trials – via outcome-dependent two-phase sampling (e.g., case-cohort) designs. These biased sampling designs have earned their popularity: they circumvent the administrative burden of collecting potentially expensive biomarker measurements on all study units without limiting opportunities to detect biomarkers mechanistically informative of the disease or infection process. To address this, we outline a semiparametric correction procedure that recovers population-level estimates, with guarantees of asymptotically efficient inference, of a causally informed vaccine efficacy measure defined by contrasting assignments of study units to active vs. control conditions while simultaneously hypothetically shifting biomarker expression in the active condition, resulting in a descriptive causal dose-response analysis informative of next-generation vaccine efficacy.

Contributed presentation 19

Almost exact Mendelian randomization

***Qingyuan Zhao**, University of Cambridge; Matt Tudball, University of Bristol; George Davey Smith, University of Bristol*

Mendelian randomization (MR) is an observational design based on the random transmission of genes from parents to offspring. However, this inferential basis is typically only implicit or used as an informal justification. As parent-offspring data becomes more widely available, we advocate a different approach to MR that is exactly based on this randomization, making explicit the common analogy between MR and a randomized controlled trial. We begin by developing a causal graphical framework for MR which formalizes several biological processes and phenomena, including population structure, gamete formation, fertilization, genetic linkage, and pleiotropy. This causal graph is then used to detect biases in the MR design and identify sufficient confounder adjustment sets to correct them. We then propose a randomization test for causal hypotheses in the MR design by using precisely the exogenous randomness in meiosis and fertilization. We term this "almost exact MR", because exactness of the inference depends on precisely knowing the distribution of offspring haplotypes resulting from meioses in one or both parents, which is widely studied in genetics. We demonstrate via simulation that propensity scores obtained from the underlying meiosis model can form powerful test statistics. Besides transparency and conceptual appeals, our approach also offers some practical advantages, including lack of commitment to a particular phenotype model, robustness to weak instruments, and eliminating bias that may arise from population structure, assortative mating, dynastic effects and linkage disequilibrium with pleiotropic variants. We conclude with a negative and positive control analysis in the Avon Longitudinal Study of Parents and Children using our R package.

Contributed presentation 20

Intergenerational effects of parental educational attainment on parenting and childhood educational outcomes: Evidence from MoBa using within family Mendelian randomization

Alexandra Havdahl, Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway; Amanda Hughes, MRC IEU at the University of Bristol; Helga Ask, Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway; Rosa Cheesman, University of Oslo, Ted Reichborn-Kjennerud, Norwegian Institute of Public Health; Ole Andreassen, University of Oslo; Elizabeth Corfield, Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway; Laurie Hannigan, Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway; Per Magnus, Norwegian Institute of Public Health, Oslo, Norway; Pål Njølstad, Univeristy of Bergen; Camilla Stoltenberg, Norwegian Institute of Public Health, Oslo, Norway; Fartein Ask Torvik, Norwegian Institute of Pub- lic Health, Oslo, Norway, Ragnhild Brandlistuen, Norwegian Institute of Public Health, Oslo, Norway; George Davey Smith, MRC IEU at the University of Bristol; Eivind Ystrom, Norwegian Institute of Public Health, Oslo, Norway; ***Neil M Davies****, Department of Statistical Sciences, University College London*

The intergenerational transmission of educational attainment is one of the most important, and studied relationships in all of social science. Longitudinal studies have found strong associations between parents and child outcomes, however, twin studies have found relatively little evidence of shared environmental effects. Here we use a novel within-family Mendelian randomization method and data from 40,910 genotyped parent-child trios from the Norwegian Mother, Father and Child Cohort (MoBa) study. We used multivariable Mendelian randomization to estimate the effects of mothers' and fathers' educational attainment controlling for the direct effects of genetic transmission. We find compelling evidence of familial effects of parental educational attainment on the educational outcomes of children aged 11 to 14 years. The within family Mendelian randomization estimates suggested that each additional year of paternal educational attainment increased their children's maths grade 9 test scores by 0.08 (95%CI:0.05,0.10)SD. Each additional year of maternal education increased their children's test scores by 0.04 (95%CI:0.01,0.07)SD. We perform pleiotropy robust sensitivity analyses using two-sample Mendelian randomization. These estimates are consistent with substantial effects on their children's educational achievement, but may be due to assortative mating or population stratification. We demonstrate that Mendelian randomization can be used to not only control for demographic factors and familial effects, but also estimate them. More research is urgently needed in this area, both to provide more samples of parent-child trios and to assess threats to inference from selection bias.

Contributed presentation 21

Bounding the selection bias

***Stina Zetterstrom**, Uppsala University; Ingeborg Waernbaum, Uppsala University*

Applying inclusion or exclusion criteria to a population defines a subpopulation after selections. The selection of the study population may induce bias for causal estimands both for the total population as well as for the selected subpopulation. In this work, we investigate a previously proposed bound for selection bias, referred to as the SV bound that is based on assumptions of values of sensitivity parameters (Smith and VanderWeele, 2019). We derive feasible regions for the sensitivity parameters of the SV bound as well as conditions for the bound to be sharp, where sharp means that the bias can be equal to the bound. In addition, we present an alternative bound that is based solely on the observed data, and is therefore referred to as the assumption free (AF) bound. For the practicing data-analyst, we provide an R package for calculating the SV and AF bounds. The package includes functions that calculate the sensitivity parameters under interpretable model assumptions and can be used for multiple selections. The theoretical properties of the bounds as well as the R package are illustrated with a simulated dataset emulating a study on the effect of zika virus on microcephaly in Brazil.

Contributed presentation 22

The built-in selection bias of hazard ratios formalized

***Richard Post**, Eindhoven University of Technology; Edwin van den Heuvel, Eindhoven University of Technology; Hein Putter, Leiden University Medical Center*

It is known that the hazard ratio lacks a useful causal interpretation. Even for data from a randomized controlled trial, the hazard ratio suffers from built-in selection bias as, over time, the individuals at risk in the exposed and unexposed are no longer exchangeable. In this work, we formalize how the observed hazard ratio evolves and deviates from the causal effect of interest in the presence of heterogeneity of the hazard rate of unexposed individuals (frailty) and heterogeneity in effect (individual modification). For the case of effect heterogeneity, we define the causal hazard ratio. We show that the expected observed hazard ratio equals the ratio of expectations of the latent variables (frailty and modifier) conditionally on survival in the world with and without exposure, respectively. Examples with gamma, inverse Gaussian and compound Poisson distributed frailty and categorical (harming, beneficial or neutral) distributed effect modifiers are presented for illustration. This set of examples shows that an observed hazard ratio with a particular value can arise for all values of the causal hazard ratio. Therefore, the hazard ratio cannot be used as a measure of the causal effect without making untestable assumptions, stressing the importance of using more appropriate estimands, such as contrasts of the survival probabilities.

Contributed presentation 23

Lifetime relative risk and applications to population attributable fractions

John Ferguson, *University of Galway; Alberto Alvarez, University of Galway; Conor Judge, University of Galway; Martin Mulligan, University of Galway; Martin O'Donnell, University of Galway*

For physiological risk factors, the age that an individual is initially exposed to the risk factor can be important in determining their risk of eventual disease. For instance, it is hypothesised that hypertension that is initially diagnosed in mid-life has a stronger causal effect on vascular dementia compared to hypertension that first appears in late life. These effects are sometimes examined in cohort studies via comparisons of exposed individuals at baseline with a 'reference' group unexposed to the risk factor at baseline at various ages. However, if the relative risk for disease depends on age at initial exposure, age-specific relative risk estimates from cohort studies will have no straightforward causal interpretation even after appropriate adjustment for confounding.

Here, we instead consider exposure to the risk factor at differing ages as differing treatments and lifetime non exposure as a reference treatment. 'Lifetime' relative risks are defined as the relative probability of eventual disease comparing potential exposure to the risk factor at differing ages to lifetime non-exposure.

Optimally, lifetime relative risks should be estimated with individual-level longitudinal data; yet often such data are unavailable. We describe an alternative procedure to approximate lifetime relative risks using summary data from published cohort studies, and detail conditions under which such estimation is valid. In addition to being of interest in their own right, lifetime relative risks are useful in estimating population attributable fractions (PAF)s. We illustrate this connection via estimation of the PAF for vascular dementia due to hypertension in the United Kingdom.

Contributed presentation 24

Unraveling exposure incidence and excess risk components of the attributable fraction of a time-dependent exposure

Johan Steen, *Department of Internal Medicine and Pediatrics, Ghent University, Renal Division and Department of Intensive Care Medicine, Ghent University Hospital; Wim Van Biesen, Department of Internal Medicine and Pediatrics, Ghent University, Renal Division, Ghent University Hospital; Johan Decruyenaere, Department of Internal Medicine and Pediatrics, Ghent University, Department of Intensive Care Medicine, Ghent University Hospital; Stijn Vansteelandt, Department of Applied Mathematics, Computer Science and Statistics, Ghent University*

The population-attributable fraction (PAF) expresses the proportion of cases by a certain time t that can be causally ascribed to a certain (time-dependent) dichotomous exposure in a given (closed) cohort. It is dependent on both (i) exposure incidence and (ii) excess risk related to exposure; two components that may change over time and that in part reflect the effectiveness of preventive and therapeutic efforts, respectively. Attempts to unravel these two components in the presence of a time-dependent exposure may be challenging and has led applied researchers to resort to approaches that target excess risk measures which lack a clear interpretation, let alone a clear causal interpretation. We propose a principled solution by extending the characterization of the PAF as a weighted version of the average causal effect of exposure in the exposed, in the case of a time-fixed exposure, to settings with time-dependent exposures, where the PAF can be shown to correspond to a weighted average of the time-dependent average causal effects of incident exposure in newly exposed patients. As a motivating example, we estimate the fraction of ICU deaths that can be attributed to incident acute kidney injury (AKI) in critical care patients, and illustrate and provide intuition into our proposal to separate the component that reflects how common AKI is from the component that reflects how harmful AKI is, using an inverse probability weighted estimator for the time-dependent effect of exposure in the exposed.

Contributed presentation 25

Is there a survival benefit of initiating treatment earlier?

***Haris Fawad**, University of Oslo; Pål Ryalen, UiO; Kjetil Røysland, UiO*

This question arises in clinical settings such as organ transplantations and palliative care for cancer patients. To provide an answer, we build a causal model based on counting processes and their intensities. The counting process framework is already well-suited for survival analysis; we use it to model hypothetical scenarios in which treatment occurs earlier. We highlight the structural assumptions needed for the identification of survival parameters, and present consistent estimators along with analytical expressions for their asymptotic confidence intervals. To illustrate, we analyse observational data from a kidney transplant study to estimate the survival effect of earlier transplantation among elderly patients.

Contributed presentation 26

Losing control (group)? The machine learning control method for counterfactual forecasting

Augusto Cerqua, Sapienza University of Rome; Marco Letta, Sapienza University of Rome; ***Fiammetta Menchetti****, DiSIA University of Florence*

The standard way of estimating treatment effects relies on the availability of a similar group of untreated units. Without it, the most widespread counterfactual methodologies cannot be applied. We tackle this limitation by presenting the Machine Learning Control Method (MLCM), a new causal inference technique for aggregate data based on counterfactual forecasting via machine learning. The MLCM is suitable for the estimation of individual, average, and conditional average treatment effects in evaluation settings with short panels and no controls. The method is formalized within the Rubin's potential outcomes framework and comes with a full set of diagnostic, performance, and placebo tests. We illustrate our methodology with an empirical application on the short-run impacts of the COVID-19 crisis on income inequality in Italy, which reveals a striking heterogeneity in the inequality effects of the pandemic across the Italian local labor markets.

Contributed presentation 27

Asymptotics of caliper matching estimators for average treatment effects

Máté Kormos, *Delft University of Technology; Stéphanie van der Pas, Amsterdam University Medical Centers; Aad van der Vaart, Delft University of Technology*

Caliper matching is used to estimate causal effects of a binary treatment from observational data by comparing matched treated and control units. Units are matched when their propensity scores, the conditional probability of receiving treatment given pretreatment covariates, are within a certain distance called caliper. So far, theoretical results on caliper matching are lacking, leaving practitioners with ad-hoc caliper choices and inference procedures. We bridge this gap by proposing a caliper that balances the quality and the number of matches. We prove that the resulting estimator of the average treatment effect, and average treatment effect on the treated, is asymptotically unbiased and normal at parametric rate. We describe the conditions under which semiparametric efficiency is obtainable, and show that when the parametric propensity score is estimated, the variance is increased for both estimands. Finally, we construct asymptotic confidence intervals for the two estimands.

Contributed presentation 28

Causal effect metrics are not equal when facing population's shifts

***Bénédicte Colnet**, INRIA; Julie Josse, INRIA; Gaël Varoquaux, INRIA; Erwan Scornet, École polytechnique*

There are currently several ways and practices to report so-called treatment or causal effect in applied work: absolute difference, ratio, odds ratio, number needed to treat, and so on. The choice of such measures have several impacts as: (1) providing different appreciation of the same phenomenon, and (2) leading to different heterogeneity of treatment effect patterns. In addition, not all metrics are collapsible (i.e. global effect can be expressed as a weighted sum of local effects). In this work, we review usual causal measures present in the literature, and recall typical arguments found about their pros and cons. Doing so, we enrich the existing formal framework and definitions of collapsibility and treatment effect heterogeneity, unifying different existing definitions. But the main contribution is our proposal to reverse the thinking. Rather than starting from the metric, we propose to start from a working model of the outcome. We show that depending on the nature of the outcome, each metric expresses something different. A by-product of our analysis is another way of understanding what heterogeneity and homogeneity of treatment effect means, not through the lens of the metric, but through the lens of the working model. Finally, we also show how some metrics are easier to generalize to other populations, in the sense that they require less covariates for standardization. Our results are general as the proposed models are non-parametric.

Contributed presentation 29

Using policy learning to reduce catastrophic health expenditures: an evaluation of Indonesia's JKN insurance schemes

***Vishalie Shah**, University of York; Taufik Hidayat, Universitas Indonesia; Andrew Jones, University of York; Noemi Kreif, University of York*

Policy learning uses observational data to learn optimal rules that maximise a policy's expected outcomes, by mapping an individual's covariate profile to a policy decision. Popular approaches involve estimating heterogeneous policy impacts via the conditional average treatment effect (CATE) function, and assigning the policy to units with a beneficial impact. Alternatives include simpler rules that may allow for easier interpretation (e.g. tree-based rules). We construct a rule that assigns households to Indonesia's two subsidised health insurance programmes, using an objective function that corresponds to the total reduction in the risk of incurring "catastrophic health expenditure"; a measure of the financial protectiveness of insurance. We use an ensemble algorithm, the super learner, that finds the best weighted combination of candidate estimators of the rule, including different regression specifications of CATE-based rules (e.g. linear models and data-adaptive models) and simpler rules (e.g. fixed-depth decision trees). We learn a second rule that constrains the proportion of households that can receive the policy, to reflect budget restrictions. We evaluate the rules using doubly robust estimators of the counterfactual mean outcomes. Our learned unconstrained and resource-constrained rules achieve better expected outcomes (i.e. a reduction in the risk of incurring catastrophic expenditures) compared to the actual assignment and static rules, that assign the same programme to all. Geography, particularly the urban-rural distinction, is the main differentiating factor between the assignment under the rules and the actual assignment, which is associated with the availability of health services and a determinant of health spending.

Contributed presentation 30

A Bayesian multivariate factor analysis model for causal inference using time series observational data on mixed outcomes

Pantelis Samartsidis, *University of Cambridge; Shaun Seaman (Cambridge) and Daniela De Angelis (Cambridge)*

Estimating the effect of an intervention from aggregate time-series observational data on multiple units and outcomes (known as policy evaluation), is a problem that arises frequently in many fields of applied research including epidemiology, econometrics and political science. Here, we propose a Bayesian factor analysis model for estimating intervention effects in such settings. Our model includes a regression component to allow adjustment for observed confounders, and a latent component to account for certain forms of unobserved confounding. Compared to existing approaches, our method is one of the few that can simultaneously deal with outcomes of mixed type (continuous, binomial, count), increase efficiency in the estimates of the causal effects by jointly modelling multiple outcomes affected by the intervention (shown via a simulation study), allow for staggered adoption of treatment, and account for uncertainty in the dimensionality of the latent component. Inference is done under the Bayesian paradigm using Markov chain Monte Carlo. Hence, the method provides uncertainty quantification for all causal estimands of interest. We further extend our approach to model effect heterogeneity and modification. Specifically, we demonstrate that modelling effect modification is not straightforward in causal factor analysis (or any method with temporal components) due to non-identifiability. We then demonstrate how this problem can be circumvented using a Bayesian modularisation approach, that prevents post-intervention data from informing a subset of the model parameters. The proposed method is used to evaluate the impact of Local Tracing Partnerships had on the effectiveness of England's Test and Trace programme for COVID-19.

Contributed presentation 31

Causal inference with continuous multiple time point interventions

***Michael Schomaker**, LMU Munich; Iván Diaz, NYU Grossman School of Medicine; Paolo Denti, University of Cape Town; Helen McIlleron, University of Cape Town*

Currently, there are limited options to estimate the treatment effect of variables that are continuous and measured at multiple time points, i.e. through the dose-response curve. However, these situations may be of relevance: in pharmacology, one may be interested in how outcomes of people living with -and treated for- HIV, such as viral failure, would vary for time-varying interventions such as different drug concentration trajectories. A challenge for doing causal inference with continuous interventions is that the positivity assumption is typically violated. To address positivity violations, we develop projection functions, which reweigh and redefine the estimand of interest based on functions of the conditional support for the respective interventions. With these functions, we obtain the desired dose-response curve in areas of enough support, and otherwise a meaningful estimand that does not require the positivity assumption. We develop g-computation type plug-in estimators for this case. Those are contrasted with using g-computation estimators in a naïve manner, i.e. applying them to continuous interventions without addressing positivity violations. The ideas are illustrated with longitudinal data from HIV+ children treated with an efavirenz-based regimen as part of the CHAPAS-3 trial, which enrolled children < 13 years in Zambia/Uganda. Simulations show in which situations a naïve g-computation approach is appropriate, and in which it leads to bias and how the proposed weighted estimation approach recovers the alternative estimand of interest.

Contributed presentation 32

Stable probability weighting: Large-Sample and finite-sample estimation and inference methods for heterogeneous causal effects of multivalued treatments under limited overlap

***Ganesh Karapakula**, University of Cambridge*

In this paper, I try to tame "Basu's elephants" (data with extreme selection on observables). I propose new practical large-sample and finite-sample methods for estimating and inferring heterogeneous causal effects (under unconfoundedness) in the empirically relevant context of limited overlap. I develop a general principle called "Stable Probability Weighting" (SPW) that can be used as an alternative to the widely used Inverse Probability Weighting (IPW) technique, which relies on strong overlap. I show that IPW (or its augmented version), when valid, is a special case of the more general SPW (or its doubly robust version), which adjusts for the extremeness of the conditional probabilities of the treatment states. The SPW principle can be implemented using several existing large-sample parametric, semiparametric, and nonparametric procedures for conditional moment models. In addition, I provide new finite-sample results that apply when unconfoundedness is plausible within fine strata. Since IPW estimation relies on the problematic reciprocal of the estimated propensity score, I develop a "Finite-Sample Stable Probability Weighting" (FPW) set-estimator that is unbiased in a sense. I also propose new finite-sample inference methods for testing a general class of weak null hypotheses. The associated computationally convenient methods, which can be used to construct valid confidence sets and to bound the finite-sample confidence distribution, are of independent interest. My large-sample and finite-sample frameworks extend to the setting of multivalued treatments. arXiv preprint: https://arxiv.org/abs/2301.05703

Contributed presentation 33

When you see nothing at all

Daniel Farewell, *Cardiff University; Rhian Daniel, Cardiff University*

If observed exposure levels do not vary within confounder strata (sometimes called a positivity violation), likelihoods may be flat. Although this can be challenging for convergence-based causal inference, we remind ourselves that there is no problem in principle: each candidate data-generating mechanism has a computable probability of giving rise to the data, and an equally calculable associated causal estimand. For select, simple such circumstances, we describe targeted causal inference that is likelihood-driven and Bayes-inspired while retaining frequentist error control.

Contributed presentation 34

Causal and counterfactual views of missing data models

Razieh Nabi, *Emory University; Rohit Bhattacharya, Williams College; Ilya Shpitser, Johns Hopkins University; James Robins, Harvard University*

It is often said that the fundamental problem of causal inference is a missing data problem – the comparison of responses to two hypothetical treatment assignments is made difficult because for every experimental unit only one potential response is observed. In this paper, we consider the implications of the converse view: that missing data problems are a form of causal inference. We make explicit how the missing data problem of recovering the complete data law from the observed law can be viewed as identification of a joint distribution over counterfactual variables corresponding to values had we (possibly contrary to fact) been able to observe them. Drawing analogies with causal inference, we show how identification assumptions in missing data can be encoded in terms of graphical models defined over counterfactual and observed variables. We review recent results in missing data identification from this viewpoint. In doing so, we note interesting similarities and differences between missing data and causal identification theories.

Contributed presentation 35

A counterfactual causal model for marked point processes

Pål C. Ryalen, *Department of Biostatistics, University of Oslo; Mats J. Stensrud, Department of Mathematics, École Polytechnique Fédérale de Lausanne; Kjetil Røysland, Department of Biostatistics, University of Oslo*

Many researchers use counterfactual variables for causal reasoning. Frameworks such as the FFRCISTG model formalizes the study of a finite number of such variables. A corresponding framework of counterfactuals in marked point process (MPP) models, which deals with events that occur in continuous time, has not yet been developed. This may be of interest in fields such as survival and event history analysis, where 1) there is a large literature of statistical methods formulated in time-continuous models, and 2) interest is frequently in effects of time-varying treatment strategies, which may often be most appropriately viewed as time-continuous processes. In this presentation, we aim to fill this gap by providing a rigorous framework for counterfactuals in MPP models. Specifically, we define weak and strong causal realizations, where strong causal realizations serve as counterfactual variables, and weak causal realizations are based on the notion of interventions on systems and certain invariances under these interventions. We show that a strong causal realization generally induces a weak causal realization, and we provide nonparametric identification results in the presence of unmeasured confounders. Additionally, we establish a correspondence between an 'invariance' assumption among factual processes and an 'exchangeability' assumption among corresponding counterfactual processes, and illustrate how these assumptions can be inferred from local independence graphs. We also describe and discuss deterministic and random treatment regimes in this setting, including regimes that depend on natural values of the treatment process. Finally, we highlight some similarities and differences with existing methods.

Contributed presentation 36

Doubly robust g-estimation of structural nested cumulative survival time models with non-ignorable, non-monotonic missing data in time-varying confounders

Yoshinori Takeuchi, *Division of Medical Statistics, Department of Social Medicine, Faculty of Medicine, Toho University, Tokyo, Japan; Sho Komukai, Department of Biomedical Statistics, Graduate School of Medicine, Osaka University, Osaka, Japan; Atsushi Goto, Department of Public Health, School of Medicine, Yokohama City University, Kanagawa, Japan; Tomohiro Shinozaki, Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan*

To examine the causal effects of time-varying treatments on survival, structural nested cumulative survival time models (SNCSTMs) are flexible and theoretically promising semiparametric models characterised by causally interpretable parameters. One concern is the prerequisite for uniformly scheduled data collection and complete data for time-varying confounders. For example, in pharmacoepidemiological studies using medical information databases, laboratory test results can be missing due to unscheduled hospital visits or non-compliance with health checkups. Furthermore, missing mechanisms data may be non-ignorable and non-monotone, invalidating the typical missing-data methods that assume ignorable or monotone missing mechanisms. We propose a novel g-estimation method for SNCSTMs with non-ignorable, non-monotonic missing data for time-varying confounders. We augment the g-estimation functions using missing probability and imputation models, incorporating a user-defined selection function, which allows sensitivity analyses to evaluate the departure of missing data from ignorable mechanisms. Using a proper selection function, our estimator is doubly robust in the sense that it is consistent if either model for missing probability or imputation of missing data is correct at each time point and if either model for propensity score or conditional expectation of counterfactual counting processes is correct. Moreover, applying frequentist-type multiple imputation yields a closed-form solution for calculating the estimator, even if time-varying confounders are missing. A simulation study evaluated our proposed method's finite sample performance and the estimator's double robustness. We also conducted sensitivity analyses in a pharmacoepidemiological study using a Japanese medical claims database, assessing the risk of hypoglycaemia in sulphonylurea-treated patients with incomplete haemoglobin A1c values.

Poster presentation 1

Investigating impacts of health policies using staggered difference-in-differences: the effects of digital triage on prescribing patterns of antibiotics

***Kate Ellis**, London School of Hygiene and Tropical Medicine*

Health policy interventions are often implemented at cluster level, where the intervention is rolled out to different units over time. Evaluation of the effects of the intervention requires consideration of which control groups and time periods to use in the analysis. Recent literature highlights that the standard two-way fixed effects difference-in-differences estimator can be biased for the ATT when roll out is staggered. Several methods have since been proposed that explicitly avoid using inappropriate units and time periods as controls. We use one such approach, proposed by Callaway and Sant'Anna (2021), to investigate the effects of adoption of a 'total digital triage' model in general practice on prescribing of antibiotics in England. The estimand of interest is the average effect of adoption for each group of practices (defined by year of adoption) in each year (ATT). We used data on 6,397 practices in England, of which 176 adopted the total digital triage system askmyGP between March 2019 and February 2021. We aggregated the ATTs for group and time period across all adopting practices, by group, and by time elapsed since adoption. We found strong evidence of a positive effect of adoption of askmyGP on antibiotic prescribing rates, albeit the effect was relatively small (ATT = 1.7 items per 1,000 patients per month, 95% CI=[1.1, 2.4]). As time since adoption increases, the effect size increases, while effects of adoption vary across groups. We used two recently proposed methods to assess the parallel trends assumption, which is intrinsic in difference-in-differences methods.

Poster presentation 2

A comparison of methods to develop individualized treatment rules on real data

Florie Brion Bouvier, *Université Paris Cité; Etienne Peyrot, Université Paris Cité; Alan Balendran, Université Paris Cité; Corentin Ségalas, Université Paris Cité; François Petit, Université Paris Cité; Raphaël Porcher, Université Paris Cité*

Identifying subgroups of patients who benefit from a treatment is a key aspect of personalized medicine. This can be achieved by developing individualized treatment rules (ITRs). Many machine learning algorithms have been proposed to develop such rules. It is unclear to what extent those algorithms lead to the same ITRs, i.e. recommending the treatment for the same individuals. To this aim, we compared the most common approaches on two randomized control trials. We distinguished two classes of algorithms to develop an ITR. The first class relies on predicting individualized treatment effects and then deriving an ITR by recommending treatment to those with predicted benefit. In the second class, algorithms directly estimate the ITR by minimizing a loss function. The majority of the algorithms we compared in this project fell under the first class. For each trial, we assessed the performance of ITRs with various metrics and by calculating the pairwise agreement between ITRs. Results showed that the ITRs obtained by the different algorithms generally had considerable disagreements regarding the individuals to be treated. A better concordance was found among algorithms of the same family. Overall, when evaluating the performance of ITRs in a hold-out validation sample, all algorithms produced ITRs with limited performance, whatever the performance in the training set, which suggests a high potential for overfitting. The methods are not interchangeable which draws some concerns on their practical use. This finding should be considered when developing ITRs in practice.

Poster presentation 3

Tree-based identification of predictive factors for non-randomized treatment comparisons

***Julia Krzykalla**, German Cancer Research Center (DKFZ); Axel Benner, DKFZ; Prof. Annette Kopp-Schneider, DKFZ*

Novel high-throughput technology provides detailed information on the biomedical characteristics of each patient's disease. These biomarkers may qualify as predictive factors that distinguish patients who benefit from a particular treatment from patients who do not. Hence, large numbers of biomarkers need to be tested in order to gain evidence for tailored treatment decisions ("personalized medicine"). Tree-based methods divide patients into subgroups with differential treatment effects in an automated and data-driven way without requiring extensive pre-specification. One approach that focuses explicitly on identifying predictive factors is the predMOB (Krzykalla et al., Stat Med 2020), a modification of model-based recursive partitioning for subgroup analyses (Seibold et al., Int J Biostat 2016). Like most other methods, the predMOB has been developed for the use on randomized data only. However, especially for rare cancers, data from clinical registries or observational studies might be the only available data source. In regard to the application to such data, we investigate the predMOB in combination with common methods for confounder adjustment (covariate adjustment, inverse probability of treatment weighting (IPTW), matching) (Krzykalla et al., arXiv 2022). Using simulation studies, we are able to show that covariate adjustment allows a correct identification of predictive factors while IPTW might also suggest confounding variables. If individual treatment effects or predictive effects are to be estimated, all investigated adjustment methods show satisfying results. For illustration, we apply the investigated versions of confounder-adjusted predMOB to the German Breast Cancer Study Group (GBSG) Trial 2.

Poster presentation 4

Median-based splitting rules for the causal tree

***Karolina Gliszczynska**, University Duisburg-Essen; Lennard Maßmann, University Duisburg-Essen*

Our paper contributes to the existing literature on tree-based methods for causal inference and treatment effect estimation in high-dimensional data. We investigate the robustness of causal trees, which are a variation of standard regression trees. Causal trees can be estimated in an honest way, i.e., that model complexity is unrestricted. The imposition of model structure and the estimation referring to a specific model structure is achieved by separate data sets through sample splitting. Our paper presents three median-based splitting rules for the causal tree, based on the Median Absolute Deviation (MAD), Least Median Square (LMS) and Median Squared Deviation (MSD). For median estimation we use the Hodges-Lehman estimator to reduce time complexity. We aim to robustify the estimation procedure with those splitting rules, especially when working with unbalanced data. We compare our method with existing ones (mean-based causal tree, transformed outcome tree, fit-based tree, squared t-statistic tree) via a simulation study. Thereby, we draw data from different distributions, flawed with a substantial amount of outliers. We also apply our approach to a real data set. In our simulation study, we further investigate improvements in coverage of confidence intervals for treatment effects. Our results show that our approach improves the robustness. The performance, however, depends on the distribution and outlier scenario. We find the MSD splitting rule to be an interesting alternative to existing splitting rules.

Poster presentation 5

On analysis of treatment effects in observational HIV/AIDS studies

***Muluneh Alene**, Department of Applied Mathematics, Computer Sciences and Statistics Ghent University, Belgium; Tadesse Awoke, Epidemiology and Biostatistics, University of Gondar, Gondar, Ethiopia; Stijn Vansteelandt, Department of Applied Mathematics, Computer Sciences and Statistics Ghent University, Belgium*

Abstract Background: Analysis of observational longitudinal studies can be difficult due to non-random treatment assignment and dependent censoring when studying the effects of treatment on time-to-event endpoints. This is common in HIV/AIDS studies, where treatments are assigned based on patient characteristics and may change over time. In this study, we evaluate the effect of Lopinavir-based treatment versus Atazanavir-based treatments on survival, using data from a multi-center institution-based retrospective follow-up study conducted in Ethiopia. Methods: We conducted a simulation study to evaluate five methods for estimating treatment-specific survival curves: inverse probability of treatment weighting (IPTW), inverse probability of treatment and censoring weighting (IPTCW), augmented inverse probability of treatment weighting (AIPTW), augmented inverse probability treatment weighting with censoring weights (AIPTW_IPCW), and augmented inverse probability of censoring weighting (AIPCW). Logistic regression and Cox proportional hazard models which includes baseline covariates, were used for the treatment, censoring, and outcome models. These methods were subsequently applied to time-to-event data from 835 patients from the Ethiopian study. Results: The simulation showed that IPTCW performed best in moderate sample sizes when all models were correctly specified, AIPCW performed best when the treatment model was mis-specified, and AIPTW_IPCW performed best when the censoring model was mis-specified. The data analysis showed that patients starting with Lopinavir-based treatments had a better survival rate than those starting with Atazanavir-based treatments. Conclusion: Overall, methods augmented with the outcome model had smaller bias and mean square error in moderate sample sizes. The study suggests that Lopinavir-based treatments are beneficial over Atazanavir-based treatments.

Poster presentation 6

Household water safety and adult health: Micro-econometric evidence from China

Jingmin Zhu, *University College London*

Inadequate safe water is a great public health issue worldwide. In China, unsafe water, inadequate sanitation, and poor hygiene accounted for 62,800 deaths and 2.81 million DALYs in 2008. During 2009- 2011, Chinese government invested in a national public health project to improve rural household water safety. To the best of our knowledge, this study is the first attempt to evaluate the health impacts of water safety improvement project in rural areas during 2009-2011 in China. Panel data from the China Health and Nutrition Survey from 1991 to 2015 are used, consisting of 67,841 adults. Difference-in-Differences estimation and Propensity Score Matching method are combined to identify a causal effect of water safety improvement on adults' health, which is significant from the perspective of both public health and economic productivity. Results suggest that rural water improvement project is effective in improving adults' health in terms of reduction in the incidence of diarrhea or stomachache. The results are robust using alternative estimation method, restricted sample, placebo test, and additional consideration of health insurance coverage expansion, which suggests that less suffering from waterborne diseases is not due to better individual health, other public health interventions or health insurance coverage. Influencing mechanism is indicated to be promoted water quality rather than water quantity. No significant health inequality associated with education and household income is detected. It implies the unneglected importance of household's access to safe water to adults' health, as well as the effectiveness of public health project for governments to invest in population's health.

Poster presentation 7

Computational twins: Next-generation causal modelling for defence

Helen Burke, Faculty; Mustafa Caglar, Faculty; ***Annabel Whipp****, Dstl*

Causal models have the potential to aid decision making, especially important where actions carry high consequences. This work undertakes applied research to demonstrate the power of encapsulating causal relationships within a Computational Twin (CT, an AI-enabled digital twin) to improve complex decision making across UK Defence and Security. This will ultimately help accelerate the move to human-machine teaming by demonstrating the value of augmenting domain expertise with the predictive power of AI techniques. We are developing a prototype in collaboration with a UK air base that will test this technology and support the shift from traditional, associative statistical models to causal models. This prototype will empower Defence planners to ask interventional and counterfactual questions where data are not available, testing the practical applications of this technology. Potential use cases for the work are diverse, and include; devising pilot training, deployment readiness, and predictive maintenance.

Thus far, organisational processes at the air base and their causal relationships have been identified and mapped. Relevant use cases have been selected through technical scoping to inform the construction of a relational process graph, sub-models, and subsequent encoding into a CT. Technical scoping obtained a detailed breakdown of the use case components, key metrics, sequence of decisions currently taken, alongside levers that will inform scenario planning powered by the CT. We will use real-world data, which will be back-tested and the AI-features iterated upon with domain experts to gauge practical applicability.

Poster presentation 8

The association between partisan and racial gerrymandering

***Christiana Drake**, University of California Davis; Xiner Zhou, University of California Davis; Bala Rajaratnam, University of California Davis*

In the United States the members of the House of Representatives are elected every 2 years directly by the voters in their respective districts. District boundaries are drawn by the States. A candidate is elected by a simple majority of the voters in the respective district. In 1965 Congress passed the Voting Rights Act which prohibits the disenfranchisement of racial minority voters. The US Supreme Court has ruled that it is not necessarily prohibited to draw boundaries in such a way as to favor one political party. This is popularly referred to as gerrymandering. However, the Supreme Court has not ruled on a case involving gerrymandering. Some states have independent or bipartisan commissions that draw boundaries, in other states the legislature draws the boundaries. It is tempting to create congressional districts in these states that favor the party that controls the legislature. It can be shown that it is possible to create districts in such a manner that a party that has more than 50% of the vote nevertheless ends up with fewer than half the congressional representatives in the state. We use the counterfactual framework to explore the relationship between partisan gerrymandering and the potential for racial disenfranchisement. In our model the counterfactual is a measure of the value of a vote in districts that would occur in a gerrymandered district vs. a district drawn by an independent commission.

Poster presentation 9

Mediating factors in the association between hearing impairments and quality of life among older people in Poland. Application of structural equation modeling

***Katarzyna Zawisza**, jagiellonian University Medical College; Michalina Gajdzica, Jagiellonian University Medical College, Faculty of Medicine, Chair of Epidemiology and Preventive Medicine, Department of Medical Sociology; Karolina Majdak, Jagiellonian University Medical College, Faculty of Medicine, Chair of Epidemiology and Preventive Medicine, Department of Medical Sociology; Beata Tobiasz-Adamczyk, Jagiellonian University Medical College, Faculty of Medicine, Chair of Epidemiology and Preventive Medicine, Department of Medical Sociology; Tomasz Grodzicki, Jagiellonian University Medical College, Faculty of Medicine, Chair of Internal Medicine and Gerontology*

The aim of the study was to applied structural equation modelling to determine whether feeling of loneliness, social networks and social participation mediate the association between hearing impairments and health-related quality of life among older people in Poland. Cross-sectional data for 1,977 older people aged 65 or older from Małopolska Region in Poland were analyzed. Face-to-face interviews were performed among randomly selected individuals from the general population. The health-related quality of life (HRQoL) was measured by the World Health Organization Quality Of Life – Age (WHOQOL–AGE) scale. The participants' hearing status was assessed based on self-reported data about the presence of hearing impairments (HI) and the daily use of hearing aids (HA). Structural equation model was conducted with Mplus Base Program v. 7.0. The results showed that among the participants two groups of people were characterized by worse quality of life in comparison to the reference group: those who did not report HI and did not use HA. The first was those who reported HI and did not use HA. In this case a significant sequential mediation effect was observed: the hearing status had an impact on social network, which in turn influenced the feeling of loneliness or social participation, affecting the quality of life. Secondly, the group of people who did not report HI but used HA also indicated worse HRQoL. In this case mediation effect of the feeling of loneliness was shown. Finally, HRQoL was not significantly lower for those who reported HI and used HA.

Poster presentation 10

Estimating and interpreting causal effects under violation of positivity

***Maria Geers**, Leibniz Institute for Prevention Research and Epidemiology – BIPS; Vanessa Didelez, University of Bremen and Leibniz Institute for Prevention Research and Epidemiology – BIPS*

A violation of the fundamental positivity assumption in causal inference leads to lack of overlap in the data and poses a challenge to the interpretation and estimation of causal effects. In principle, the treatment strategies should only be compared for those units for whom both strategies are possible and sensible. In a target trial emulation (TTE) this would be ensured by carefully chosen eligibility criteria. Some methods of estimation provide a more or less automated approach to lack of overlap, e.g. double machine learning or overlap weighting; moreover, propensity score (PS) matching addresses the problem by pruning unmatched observations. Often it is not clearly emphasized that these different approaches implicitly modify the estimand and/or the population. Another difficulty concerns diagnostics for lack of overlap which often involve a subjective assessment of PS plots (which in turn depend on the chosen model for the PS). In this work, we provide a general comparison of approaches and methods of estimation for causal effects regarding the estimands, the population and the handling of positivity violations; we further review and compare techniques to diagnose lack of overlap. We illustrate this with semi-simulated data using the Rotterdam breast cancer dataset to estimate the causal effect of a hormonal therapy after a breast cancer diagnosis. Our results provide guidance on the strengths and limitations of the different methods in practical applications. Finally, using a TTE framework, we discuss what aspects of causal effect estimation under lack of positivity may or may not lend themselves to an automatization and where expert input is required.

Poster presentation 11

Identifying the marginal effects of specific educational decisions through genetic instrumental variables

***Eleanor Sanderson**, University of Bristol; Neil Davies, UCL*

On average, people who remain in education for longer have better health and socioeconomic outcomes later in life. The effects of educational attainment on many outcomes are likely to be heterogeneous with some levels of achievement having a greater impact on outcomes than others, however unmeasured confounding biases the observed associations. Instrumental variables (IV) analysis can be used to estimate causal effects in the presence of confounding if specific assumptions hold. A range of IVs have been used to test whether educational attainment causally affects health and socioeconomic outcomes. However, most of these studies assume either that each additional year of education has a constant effect or considers only the effect of obtaining a particular level of educational attainment. We propose an alternative approach that uses genetic variants as IV's to estimate the marginal causal effect of different estimation decision points through multivariable IV estimation. Genetic variants associated with educational attainment are potentially valid IVs as they are randomly inherited from parents at conception, and cannot be affected by factors that occur after conception. If these genetic variants have slightly different effects on each stage of education they can be used to identify the marginal effects of different stages of educational attainment. Here, we describe the approach, its assumptions, limitations, and data requirements. We illustrate the approach using data from the UK Biobank to estimate the marginal causal effect of remaining in school to age 16, 18, and attending university on a set of health and socioeconomic outcomes.

Poster presentation 12

Recoverability of causal effects in a longitudinal study using missingness DAGs

Anastasiia Holovchak, *Department of Statistics, Ludwig Maximilian University of Munich; Michael Schomaker, Department of Statistics, Ludwig Maximilian University of Munich; Paolo Denti, UCT Pharmacometrics, University of Cape Town; Helen McIlleron, UCT Pharmacometrics, University of Cape Town*

Missing data in multiple variables is a common issue. We investigate the applicability of the framework of graphical models for handling missing data to a complex longitudinal pharmacological study of HIV-positive children treated with an efavirenz-based regimen as part of the CHAPAS-3 trial. Specifically, we examine whether the causal effects of interest, defined through static interventions on multiple continuous variables, can be recovered (estimated consistently) from the available data only. So far, there exists no general algorithm for deciding on recoverability, and decisions have to be made on a case-by-case basis. We emphasize sensitivity of recoverability to even the smallest changes in the graph structure, and present recoverability results for three plausible missingness DAGs in the CHAPAS-3 study (directed acyclic graphs), informed by clinical knowledge. Further, we propose the concept of "closed missingness mechanisms" and show that under these mechanisms an available case analysis is admissible for consistent estimation for any type of statistical and causal query, even if the underlying missingness mechanism is of MNAR type. Simulations demonstrate how estimation results vary depending on the modelled missingness DAG. Our analyses are possibly the first to show the applicability of missingness DAGs to complex longitudinal real-world data, while highlighting the sensitivity with respect to the assumed causal model.

Poster presentation 13

Fast machine learning causal network analysis using genetic instruments

Hui Guo, *Centre for Biostatistics, University of Manchester; Kelly Jack, Centre for Biostatistics, University of Manchester; Berzuini Carlo, Centre for Biostatistics, University of Manchester*

Mendelian randomization (MR) is a popular method of utilising genetic variants as instruments to investigate causal relationships between risk factors and disease outcomes in observational studies. It mimics randomized controlled trials by assuming that each individual's genes were inherited randomly from their parents. MR sidesteps the issue of confounding (both observed and unobserved). There has been a large growth in the application of MR approaches to health research. Most MR methods require parametric models. When sample size and/or the number of variables increases, it is likely that this approach becomes computationally intractable.

Network analysis, mainly using machine learning algorithms (e.g. Bayesian network), allows for inclusion of many risk factors and a disease outcome simultaneously in a single model, with the aim to identify direct and indirect effects of the risk factors on the outcome. An advantage of this approach is that it does not require a pre-specified parametric model or impose restrictive assumptions. Hence, they can potentially reduce bias compared to mis-specified complex models. Network analysis is an effective way of exploring data structure by discarding redundant variables (e.g., the PC algorithm) after testing for marginal and conditional (in)dependence properties between each pair of variables. However, the relationships these networks depict are often associations/correlations, and thus, do not necessarily have a causal interpretation.

I will discuss how one can take forward strengths of both MR and machine learning for causal network analysis. I will also introduce a way of improving computational efficiency in causal network analysis (work in progress).

Poster presentation 14

Quasi-experimental methods with pooled infectious disease data

***Heather Hufstedler**, University Hospital, Heidelberg*

Causal inference methods, such as instrumental variables (IV), regression discontinuity (RD), difference-in-difference (DiD), or marginal structural models (MSM), e.g., are potentially much more effective at dealing with time-varying confounding and controlling for measured and unmeasured confounding than standard statistical methods or regression-based adjustments. Unfortunately, according to our findings in two recent systematic reviews, such causal inference methods are not widely implemented by infectious disease researchers who pool individual-level data, despite their utility. We apply some of these methods to newly assembled pooled individual-level patient data collected during the 2014-2016 Ebola outbreak in West Africa, and discuss challenges and opportunities that the data presents. The focus of our investigation is on highlight viable sources of exogenous variation for identifying the causal effect of treatment on patient survival.

Poster presentation 15

Causal inference approaches using multi-level treatments in medical research, a systematic review

***François Bettega**, HP2; Clémence Leyrat, LSHTM; Sébastien Bailly, HP2*

Causal inference methods based on observational data represent an alternative to randomised controlled trials (RCT) when RCTs are not feasible. Inverse-probability-of-treatment weighting (IPTW) is one of the most popular of these methods to address confounding. In medical research, IPTW is mainly applied to binary treatments even when multiple treatments are considered, despite the availability of IPTW estimators for multiple treatment levels. We conducted a systematic review of medical publications using IPTW methods for multilevel treatments. Our objective was to assess how these methods are applied in practice, and the quality of their reporting. From 5299 screened article, 91 – from 17 medical specialties -were retained in the final analysis. The number of treatment groups varied between 3 and 11 with a large majority of articles (72 (79.1%)) with 3 or 4 groups. The most commonly used method for estimating the weights was multinomial regression (59 (64.8%)) and generalized boosted models (26 (28.6%)). The covariates of the weighting model are presented in 13 articles (14.3 %), 52 articles (57.1 %) did not discuss the covariate balance after weighting and few articles refer to the assumptions needed to obtain correct inferences 9 (9.8 %). This literature review shows that IPTW with multilevel treatment could be applied more frequently in the medical literature and its reporting is suboptimal.

Poster presentation 16

Nonparametric Mendelian randomisation for characterising nonlinear exposure-outcome relationship with discrete instrumental variables

***Cunhao Liu**, MRC Biostatistics Unit, University of Cambridge, Stephen Burgess, MRC Biostatistics Unit and Cardiovascular Epidemiology Unit, University of Cambridge*

In many practical situations, researchers are interested in understanding nonlinear causal effects of an exposure on an outcome. Nonlinear Mendelian randomisation (MR) is an extension to standard MR to characterise potentially nonlinear exposure-outcome relationships using genetic variants as instrumental variables (IV). The approach divides the population into strata with different average levels of the exposure, and estimates average causal effect in each stratum, known as the localised average causal effect (LACE). However, this method typically requires a strong constant genetic effect assumption for the exposure-instrument relationship, and is unable to capture complex nonlinearities, for example a threshold relationship. We propose a nonparametric extension to nonlinear MR with discrete instrumental variables, which we call nonparametric MR, that does not rely on the constant genetic effect assumption and is able to flexibly characterise much more complex nonlinear relationships. We argue that our nonparametric MR estimates converge to a different version of localised average causal effect, which we call quantile average causal effect (QACE), that approximates the derivative of the true exposure-outcome function. Our method works well when the instrument is weak and takes only two or three values, a situation which most of the existing nonparametric IV methods struggle to deal with. Our simulation study shows that nonparametric MR consistently outperforms nonlinear MR under various scenarios and is robust to very weak instruments. We also illustrate our method using a real data example from UK Biobank, showing a nonlinear causal relationship between BMI and pulse rate.

Poster presentation 17

Mediation analysis to reveal the causal relationship between sex and in-hospital mortality in patients with acute myocardial infarction: Real-world findings from the Lombardy health database

Alice Bonomi, *IRCCS Centro Cardiologico Monzino, Milan, Italy; Arianna Galotta, IRCCS Centro Cardiologico Monzino, Milan, Italy; Erica Rurali,IRCCS Centro Cardiologico Monzino, Milan, Italy; Nicola Cosentino, IRCCS Centro Cardiologico Monzino, Milan, Italy;Giancarlo Marenzi, IRCCS Centro Cardiologico Monzino, Milan, Italy*

In acute myocardial infarction (AMI), women have a higher in-hospital mortality compared with men. This can be due to older age and multimorbidity. We investigated whether this is true in a large population of AMI patients by using mediation analysis. We analyzed 265,048 patients (35% women) with AMI in the period 2003-2018 from administrative Lombardy Health database (Italy). The primary outcome was in-hospital mortality. The mediation analysis was performed by SAS Proc CALIS and Proc CAUSALMED procedure. The strength of direct and indirect relationships between variables was quantified by standardized beta coefficients. We implemented two statistical models: 1) with only one mediator (age); 2) with the following possible mediators: age, diabetes mellitus (DM), atrial fibrillation (FA), chronic obstructive pulmonary disease (COPD) and percutaneous coronary intervention (PCI) use during index hospitalization. In the first model, the gender had both direct and indirect effects on in-hospital mortality (Beta=0.018±0.002 and 0.067±0.001, respectively). All effects and their standard error estimates closely matched those PROC CAUSALMED. In the second model, only PROC CALIS could be implemented; the direct and indirect effects of gender on in-hospital mortality were 10.6% (Beta=0.008±0.002) and 89.4% (Beta=0.076±0.001), respectively. Specifically, indirect effect passed through 54% age, 32% PCI, 8% COPD, 4% FA and 2% DM. In conclusion, our data showed that the higher in-hospital mortality of women with AMI is possibly caused by their higher prevalence of comorbidities. The implementation of Proc CALIS and Proc CAUSALMED is indifferent; however, only Proc Proc CALIS can be implemented in case of multiple mediators.

Poster presentation 18

How effective are lockdowns at reducing infection numbers? Evidence from a Danish quasi-experiment

Florian Ege, *University of Southern Denmark (SDU)*

In late 2020 Denmark became the focus of the global efforts against the Covid-19 pandemic when reports of a mutated Covid-19 virus variant in the country's 17 million strong mink population appeared to threaten efforts to develop an effective vaccine. In this study the ensuing short-lived yet stringent lockdowns imposed in 7 of the country's 98 municipalities as a measure to halt the spread of the new virus variant among humans are analyzed for their effectiveness. Synthetic counterfactuals are created for each of the 7 municipalities using a weighted average combination of the the 91 municipalities not targeted by the stringent measures. This allows for a clear overview regarding the development of daily infection numbers in the municipalities as well as their counterfactuals. Hereafter, reductions in mobility within each of the 7 municipalities are analyzed to better understand the implementation, timing and variations of the stringent lockdowns policies. It becomes evident that in the vast majority of the 7 municipalities, the short-lived, stringent lockdowns were able to abruptly halt additional infections, indicating a substantial and significant reduction in daily infections when compared to the respective synthetic counterfactual municipality.

Poster presentation 19

SMaC: Spatial matrix completion method

Giulio Grossi, *University of Florence; Alessandra Mattei, University of Florence; Georgia Papadogeorgou, University of Florida*

Synthetic control methods are commonly used in panel data settings to evaluate the effect of an intervention. In many of these cases, the treated and control time series correspond to spatial areas such as regions or neighborhoods. We work in the setting where a treatment is applied at a given location and its effect can emanate across space. Then, an area of certain size around the intervention point is considered to be the treated area. Synthetic control methods can be used to evaluate the effect that the treatment had in the treated area, but it is often unclear how far the treatment's effect propagates. Therefore, researchers might consider treated areas of different sizes and apply synthetic control methods separately for each one of them. However, this approach ignores the spatial structure of the data, and can lead to efficiency loss in spatial settings. We propose to deal with these issues by developing a Bayesian spatial matrix completion framework that allows us to predict the missing potential outcomes in the different areas around the intervention point while accounting for the spatial structure of the data. The missing time series in the absence of treatment for the treated areas of all sizes are imputed using a weighted average of control time series, where the weights are assumed to vary smoothly over space according to a Gaussian process. Our motivating application is the construction of the first line of the Florentine tramway, which affect business presence at various distances from the tramway stops.

Poster presentation 20

Super efficient estimation for a sieve of statistical models

***Ivana Malenica**, Harvard University*

There is an increasing interest in identifying and estimating causal effects in longitudinal settings, where dependence is prevalent across time and/or samples. In order to make progress, assumptions on the statistical model are common: usually in the form of assumed known network or Markov order, or conditional independence given (unknown) fixed dimensional summary of the past. In this work, we propose a sieve-based approach for data-adaptively picking a statistical model from an initial fully nonparametric space. We consider the case that we observe a random variable on t time-points (or n samples) with a probability distribution belonging to an infinite dimensional statistical model. We are concerned with statistical inference for a particular pathwise differentiable target parameter of the data distribution, where there is enough independence that the canonical gradient is asymptotically normally distributed. One particular example considered is estimation of the causal effect of an intervention at time t on the proximal outcome given the past, averaged over all observed times. We consider a sequence of nested models indexed by a multivariate parameter that approximates the actual statistical model as each component reaches it maximal allowed value. We propose a data-adaptive selector of this multivariate index, and estimate the target parameter with the corresponding targeted minimum loss estimator (TMLE). We establish that, under regularity conditions, this adaptive TMLE is asymptotically normally distributed and super-efficient. This estimation approach is viewed as an important alternative to the TMLE for the actual statistical model when it is believed that the model is too large to be informative, but can be captured by a smaller model in a sieve, as is often the case in structured dependent settings.

Poster presentation 21

A causal mediation analysis for a non-rare dichotomous outcome with dichotomous mediator and exposure-mediator interaction: preterm birth and maternal distress with child disability as a mediator

Laura Montelisciani, *Bicocca University of Milan; Francesca, Graziano, Bicocca University of Milan ; Marina, Cuttini, Bambino Gesù Children's Hospital*

Background/Aims: The literature on the relationship between preterm birth and maternal distress is inconsistent, with some studies finding higher rates in mothers of preterm compared to term infants while others showed no difference. We assessed the causal effect of gestational age (GA) on maternal distress and the possible mediating role of child disability. Methods: The data on 733 mothers of preterm (GA<32 weeks) and 380 concurrent term-born controls were analysed. To investigate the mediation role of child disability, a causal mediation analysis was applied. According to identifiability assumptions for a causal interpretation, maternal age, education and twins were considered confounders. The Average Causal Mediation Effect (ACME) and the Average Direct Effect (ADE) were assessed, both in terms of relative risk (RR). The RR was obtained by Causal Mediation Analysis based on two modified Poisson regression models. Additionally, the interaction between exposure and mediator was considered in the model. Results: Only a significant modest ACME was present between preterm delivery and maternal distress (RR=1.02, p=0.030), but a remarkable exposure-mediator interaction was found with a significantly increased risk of maternal distress in mothers of disabled children born at term compared to those with no disability (RR=1.70, p=0.005). Conclusion: The effect of preterm birth on maternal health, although completely mediated by the child disability , is not particularly relevant in the whole sample. The impact is stronger only in full-term mothers. In contrast with preterm birth, child disability may be completely unexpected during a full-term pregnancy, possibly increasing the likelihood of maternal distress.

Poster presentation 22

Going beyond causal strength: A new metric for measuring specificity in causal relationships

***Francisco Simoes**, Utrecht University*

The average causal effect (ACE) is a widely used metric for characterizing the strength of a causal relationship. However, causal strength is but one aspect of a causal relationship and there are many other properties that are important to consider when studying a causal relationship. One such property is specificity, which is the degree to which an intervention specifies an outcome, and vice-versa. Recently, information theoretical measures of specificity have been proposed in the philosophy of science literature, including a causal version of mutual information. However, these measures are mostly implicitly discussed and lack concrete definitions. More importantly, these measures alone are, I argue, not sufficient to capture specificity. I define these quantities concretely, and propose a new metric called reverse causal conditional entropy, which when used in conjunction with causal mutual information, enables the measurement of specificity. Additionally, I highlight how these information theoretical measures are dependent on certain arbitrary choices and how different choices lead to different flavors of specificity. My work provides insight into how these choices are made and includes illustrative examples to clarify these concepts. This research lays the foundation for the development of specificity estimators that, when used in conjunction with ACE, will provide a more comprehensive understanding of the nature and utility of a given causal relationship.

Poster presentation 23

Population-adjusted indirect comparisons: A methodological review

***Arnaud Serret-Larmande**, Université Paris Cité; Belkacem Zenati, Sorbonne University; Agnès Dechartes, Sorbonne University; Jérôme Lambert, Université Paris Cité; David Hajage, Sorbonne University*

Introduction Population-Adjusted Indirect Comparisons (PAIC) have been developed in the 2010s to allow comparison between two treatments evaluated in different trials while accounting for differences in patient characteristics, when individual patient data (IPD) are only available for one trial. They are typically used in market authorization applications where a pharmaceutical company compares its new treatment, with IPD, to another treatment developed by a competitor, with only aggregated baseline characteristics and outcomes results available in the study publication. This study aimed to describe the characteristics of these articles, evaluate the quality of their methodology and its reporting, and describe reported results. Methods Original articles reporting the use of PAIC were searched on PubMed between January 1, 2010, and April 2, 2022. Article selection and data extraction were performed independently by two reviewers Results One hundred thirty-four publications were included, which reported the results of 289 indirect comparisons. Seventy-three (55%) articles were published since January 1, 2020, and 72 (54%) pertained to onco-hematology. Pharmaceutical industry was involved in 131 (98%) articles. Key methodological aspects were inconsistently reported, e.g. 104 (36%) indirect comparisons did not report how adjustment variables were selected. Efficacy was statistically in favor of the treatment evaluated on IPD for 159 comparisons (55%), not significant in 129 (45%) and in favor of the treatment evaluated on aggregated data for only 1 comparison. Conclusion Although the number of published PAIC is rising, quality of methodology and reporting needs to be improved. Moreover, our study strongly suggests a major publication bias.

Poster presentation 24

Exploring the link between physical and mental health through the comparison of several causal inference methodologies in application to a large observational dataset

Kristina Krieger, *Lancaster University; Jo Knight, Lancaster University; Anastasia Ushakova, Lancaster University*

There is emerging evidence which suggests that there is a link between mental and physical health. Using a large observational dataset, in this ongoing project, we aim to determine and evaluate such association using a variety of causal inference methodologies. As part of this investigation, we aim to understand how the relationship between physical health for individuals with terminal illness and their mental health may have further been affected by the circumstances surrounding the COVID-19 pandemic. We will apply and overview several established causal inference methods on a large and representative of UK population longitudinal dataset, Understanding Society. This dataset provides rich information on individual socio-economic characteristics, physical health, as well as mental health. The methods which we will evaluate for ease of implementation and results interpretation, are Differences-in-Differences (DID) with multiple time periods and groups, Bayesian Additive Regression Trees (BART), Longitudinal Targeted Maximum Likelihood Estimation (LTMLE), and Structural Causal Models (SCM). Such detailed and comprehensive methods comparison will have a potential to inform the development and application of causal methods for observational data, as well as highlight methodological frameworks which have not been explored in a health and social science domain.

Poster presentation 25

Comparison between inverse probability weight and instrumental variable approaches to evaluate the effect of ICP monitoring on unfavorable outcome in a multicentre cohort study

Francesca Graziano, *School of Medicine and Surgery, Bicocca Center of Bioinformatics, Biostatistics and Bioimaging, University of Milano-Bicocca, Monza, Italy; Laura Montelisciani, School of Medicine and Surgery, Bicocca Center of Bioinformatics, Biostatistics and Bioimaging, University of Milano-Bicocca, Monza, Italy; Paola Rebora, School of Medicine and Surgery, Bicocca Center of Bioinformatics, Biostatistics and Bioimaging University of Milano-Bicocca, Monza, Italy; Stefania Galimberti, School of Medicine and Surgery, Bicocca Center of Bioinformatics, Biostatistics and Bioimaging University of Milano-Bicocca, Monza, Italy; Giuseppe Citerio, Neuroscience Department, NeuroIntensive Care Unit, Hospital San Gerardo, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy.*

The benefits of intracranial pressure (ICP) monitoring in patients with aneurysmal subarachnoid hemorrhage (SAH) on outcomes are uncertain. The greatest issue with trying to answer such a question within an observational study is confounding by indication. A relatively novel method to adjust for confounding is instrumental variable (IV) analysis. In IV analysis, a substitute variable, the instrument, is used as level of analysis. The present study aims to estimate the association of ICP monitoring (ICPm) with unfavorable outcome comparing the results from propensity score weighting (IPW) and IV analysis. 423 SAH patients from a large, international, multicentre observational study (SYNAPSE-ICU) admitted in 92 hospitals were included in this subanalysis. In the IPW approach, pseudo-populations were created using inverse probability of ICPm weights computed by a logistic regression model in which the treatment assignment was regressed on relevant covariates. In the IV approach, hospital was identified as the IV and it was calculated as the tendency per each hospital to insert ICPm adjusted by relevant covariates. In severe patients, the ICPm might reduce unfavorable outcome (using IPW approach, Odds Ratio, OR=0.14, 95%CI=0.02–0.53). The tendency of the beneficial effect is confirmed using IV analysis (OR=0.83, 95%CI=0.63-0.99). Our results suggest that the effect estimation strongly depends on the analytical method used. In the presence of unobserved confounders and expected relation between treatment and performance of the hospital, IV analysis provides valid estimates. Further studies are needed to understand the roles of different approaches in the comparative observational efficacy research.

Poster presentation 26

Assessing links between gut microbiota and diarrhea in lambs

Laurianne Voland, *Université Clermont Auvergne; Dominique Graviou, Université Clermont Auvergne; Karine Vazeille, Unité Expérimentale Herbipôle; Diego P. Morgavi, Université Clermont Auvergne; Milka Popova, Université Clermont Auvergne*

In many sheep production systems, artificial rearing of newborn lambs is common. However, this separation from the dam affects the gastrointestinal microbiota establishment, leading to higher occurrences of diarrhea and high lamb mortality with negative consequences on welfare and economic outcomes.

We designed a sequential 2-step study to determine the cause-effect relationship between gut microbiota colonisation and diarrhea occurrences. In the 1st trial, triplet lambs were allocated into three groups, M: lambs reared with their dams, AA: reared artificially, no contact with adults, AAC: AA with some contacts. We aimed to meet two conditions to establish causality: time order criterion (with repeated sampling covering the time before and after diarrhea occurrence) and the non-spurious relationship (working with 3 groups with expected differences in the colonisation process). The 2nd trial design (planned for this spring) with extreme groups alone (M and AA) would eliminate alternative causes (meeting the third condition to establish causality).

Initial results showed that the average daily gain in M was greater than in other groups, suggesting a different performance linked to rearing conditions. In contrast, diarrhea was more frequent in AA and AAC groups. However, diarrhea was also recorded in a few M lambs. These latter results will be used to assess if the microbial indicators of diarrhea differ or not between groups.

Further work is to analyse fermentation end-products from rumen contents and sequencing of rumen/feces DNA and RNA. A Bayesian model will be used for assessing causality links.

Poster presentation 27

BART as a Gaussian process

Giacomo Petrillo, *University of Florence (UNIFI), Department of Statistics, Computer Science and Applications (DiSIA)*

BART is a nonparametric Bayesian regression technique, competitive with other black-box methods. In a certain limit, it becomes equivalent to Bayesian Gaussian process (GP) regression. We exploit this connection, both theoretically and practically. We derive a computationally viable form of the covariance function of BART, which we use to characterize and interpret the prior distribution, and to make inference with its GP limit, which, contrary to the original, allows to write an analytical posterior distribution, which simplifies model building/variation, relevant for applications in causal inference (e.g., the Bayesian causal forest method of Hanh et al. 2020), as highlighted by the ACIC Competition (Dorie et al. 2019). In some conditions, it also leads to faster computation.

Although the literature currently suggests that the GP limit is not profitable in practice (Linero 2017), we investigate this thoroughly and find otherwise. We conjecture that the approximation previously used for the covariance function of BART is not close enough to the accurate covariance function we derived. On 50% of the 42 datasets of Chipman et al. (2010), the test set prediction RMSE is the same for standard BART and its GP limit. On the other half, more often than not the GP limit is worse, with $<\tilde{}10\%$ higher RMSE. This small difference shows that the GP formulation is a promising direction of model research. These results are at fixed hyperparameters; we expect the GP formulation to fare better with hyperparameter optimization. We plan to assess the method on past ACIC competitions data.

Poster presentation 28

Inferring the effect of a randomised treatment on a recurrent event process under dependent censoring

***Wout Waterschoot**, Ghent University; Stijn Vansteelandt, Ghent University; Andrea Callegaro, GlaxoSmithKline Pharmaceuticals NV*

This work is motivated by a recent GSK-guided clinical trial (NCT03281876, November 2017- March 2020) to evaluate the efficacy of a randomised NTHi-Mcat vaccine versus placebo against recurrent severe exacerbations in patients suffering from acute exacerbations of chronic obstructive pulmonary disease (AECOPD). The withdrawal of individuals throughout the course of the study complicates vaccine efficacy estimation. In particular, the exacerbation history was found to be strongly predictive for both time to withdrawal and time to future exacerbations, making naïve analyses vulnerable to selection bias. In this talk, I will present three different methods that consistently estimate the ratio of the expected number of counts experienced by a given time point in treated versus untreated individuals under the assumption that the time to censoring and the time to future exacerbations are independent conditional on the measured covariate history. The first approach comes down to an inverse probability of censoring weighted (IPCW) regression method for count data, but is inefficient by considering just a single time. We therefore next introduce an IPCW-GEE variant which considers multiple time points, but is found computationally inefficient. To accommodate these concerns, we propose a novel weighting technique, which we call hazard inverse probability of censoring weighting (HIPCW). It retains the simplicity of IPCW, but extracts more efficiency by using the last recorded outcome of each individual. The proposed methods are validated by extensive simulations and through re-analysis of clinical trial NCT03281876.

Poster presentation 29

A comparison of calibration and entropy balancing estimators of the average causal effect

David Källberg, *Umeå School of Business, Economics and Statistics (USBE), Umeå University, Umeå, Sweden; Ingeborg Waernbaum, Department of Statistics, Uppsala University, Sweden, and Institute for Evaluation of Labour Market and Education Policy, IFAU, Uppsala, Sweden; Emma Persson, Umeå School of Business, Economics and Statistics (USBE), Umeå University, Umeå, Sweden*

Standard inverse probability weighting estimators of average causal effects that build on a propensity score model can be unstable and fail to balance the observed covariates. Entropy balancing, or calibration, is an alternative class of weighting estimators that by construction satisfy balance constraints and stability properties. The calibrated weights are obtained by minimizing a measure of dispersion while the constraints are met for the selected covariate functions. In this paper, we study and compare a class of calibration estimators for the average causal effect of a binary treatment. The estimators can be divided into two groups defined by the type of balance constraints i) two-way balance between the treatment and control group, and ii) three-way balance between the treated, control, and combined group. Large sample properties of the estimators including limiting estimation error, consistency and asymptotic normality, are studied and compared under the assumption of balancing a fixed set of constraints. Corresponding variance estimators are also proposed. We show implications of two- and three-way balance, respectively. The estimators are compared in a simulation study and applied in a motivating example estimating the effect of low school grades on acute complications of Type 1 Diabetes mellitus.

Poster presentation 30

Bayesian multiple bias modeling for unobserved confounding and measurement error: An application to the effect of ADHD medication on school performance in Norwegian children diagnosed with ADHD

Tomás Varnet Pérez, *Norwegian Institute of Public Health (Department of Child Health and Development), University of Oslo (Faculty of Medicine); Guido Biele, Norwegian Institute of Public Health (Department of Child Health and Development)*

Unobserved confounding and measurement error are considered some of the main validity threats to observational research. Conventional sensitivity analyses that focus on a single bias source at a time can be overly optimistic about the extent of bias present. On the other hand, sensitivity analyses based on bounds can be overly pessimistic when multiple bias sources are at play. Thus, our work in progress uses Bayesian multiple sensitivity analysis to better take into account the uncertainty from multiple bias sources on effect estimation. More specifically, we implement a joint model that considers (a) synthetic data to model an unobserved confounder, (b) a mixture model for the measurement error and (c) the possible interaction of both. We present simulation results showing that, if we know the true data generating process, we can recover an unbiased estimate of the effect of an intervention. We also present results from an application to the estimation of the causal effect of ADHD medication on the school performance of Norwegian children diagnosed with ADHD. The example data comes from Norwegian registries, covering children cohorts born between 2000 and 2007, for which we know there is both unobserved confounding (due to unmeasured variables such as family conflict) and measurement error (due to noncompliance of pharmaceutical treatment).

Poster presentation 31

dpa: a flexible R package implementation of dynamic path analysis for longitudinal mediation analysis of time-to-event data with application to a phase III clinical cancer trial

Matthias Kormaksson, Novartis; David Demanse, Novartis; Markus Lange, Novartis; Jiawei Duan, Novartis; Qing Xie, Novartis; Mariana Carbini, Novartis; Claudia Bossen, Novartis; ***Susanne Strohmaier****, Medical University of Vienna; Achim Guettner, Novartis; Antonella Maniero, Novartis*

Why does a beneficial treatment effect on a longitudinal biomarker not translate into overall treatment benefit on survival, when the biomarker is in fact a prognostic factor of survival? In a recent exploratory data analysis in oncology, we were faced with this seemingly paradoxical result. To address this problem, we applied Dynamic Path Analysis (DPA), which allows us to perform mediation analysis with a longitudinal mediator and a time-to-event outcome. More specifically, modeling is carried out via dynamic path models, involving the additive model for the hazard and recursive linear least squares regressions for the mediator model to estimate cumulative direct, indirect and total treatment effects.

The additive model has been implemented in the "timereg" R-package on CRAN while a user-friendly implementation of the estimating equations for the DPA framework is not publicly accessible. Due to the relevance of the methodology in drug development we implemented the modeling and estimating procedures in an R-package called "dpa" along with plotting routines for visualizing the direct, indirect, and total effects.

In this talk we will present the "dpa" R-package and demonstrate its functionalities on data from a phase III clinical cancer trial aiming to shed light on the underlying mechanisms that could explain our initial results. We will also discuss how our research question is compatible with the framework of separable effects suggested by Didelez (2018) and hence the potential causal interpretation of the estimated direct and indirect effects.

Poster presentation 32

Stereotactic ablative radiotherapy versus video-assisted lobectomy for operable stage I non-small-cell lung cancer: Study protocol for an emulated target trial

***Ahmed Bedir**, Department of Radiation Oncology, University Hospital Halle (Saale), Germany; Lamiaa Hassan, Institute of Medical Epidemiology, Biometry, and Informatics Martin Luther University Halle-Wittenberg, Germany; Prof. Dirk Vordermark, Department of Radiation Oncology University Hospital Halle (Saale), Germany; PD Dr. Daniel Medenwald, Department of Radiation Oncology University Hospital Halle (Saale), Germany.*

Introduction: Video-assisted lobectomy (VATS) is a commonly employed surgical technique for the management of operable early stage non-small cell lung cancer (NSCLC) tumors. However, the applicability of this procedure is dependent upon the patient's ability to tolerate surgery. In light of this, stereotactic body radiotherapy (SABR) has emerged as a viable alternative treatment strategy for patients who are inoperable or who refuse surgery. Considering the lack of randomized controlled trials and the inherent biases in observational cohort studies, this study protocol proposes an emulated target trial design to investigate the causative effect of SABR, in comparison to VATS, on overall survival in operable early stage NSCLC patients.

Materials and analysis: Data on NSCLC patients will be collected from secondary care administrative records, linked to German cancer registry data. The protocol was developed using the target trial methodology outlined by Hernan et al. It establishes specific parameters for the components in order to mitigate bias in the analysis of observational data and facilitate the calculation of causal estimands.The target trial design that would be emulated, is a multicenter open-label two-parallel arm superiority randomized trial. Mediators and confounding variables were determined through the use of a directed acyclic graph. The statistical analysis aims to measure the per-protocol effect of SABR versus VATS within three months of diagnosis on survival, through the difference in restricted mean survival times, using weighted nonparametric Kaplan-Meier curves.

Poster presentation 33

Interactive causal discovery and inference for comorbidity onset and progression in SCI population

***Yanke Li**, ETH Zurich; Anke Scheel-Sailer, Swiss Paraplegic Centre; Robert Riener, ETH Zurich & University of Zurich; Diego Paez-Granados, ETH Zurich & Swiss Paraplegic Centre*

Spinal Cord Injury (SCI) is a lifelong chronic condition affecting many functions and systems in the body. Despite extensive research on risk factors towards SCI comorbidity onset, the underlying causal relations and relevant biomarkers for early detection remain unclear, which would help medical workers make timely diagnoses, prevention, and interventions. Therefore, within the framework of continuous health monitoring in long-term chronic patients, we aspire to build a general pipeline for the intelligent assistive causal learning system based on clinical and sensory-systems data combining expert knowledge for continual causal discovery and inference in disease onset and progression. The human-machine interactive causal discovery with the consequent theoretical guarantee is this project's focus, with a novel application on multi-modal data including both static descriptive features extracted retrospectively from patient medical records and dynamic sensor signals collected from subjects under controlled interventions. We will present an iterative approach to learning the directed maximal ancestral graph, where an expert's belief is embedded into the constrained-based causal discovery algorithm by adjusting the corresponding rejection threshold in conditional independence tests. In this way, we aim to improve causal discovery with expert knowledge integration that simultaneously accounts for the uncertainty of the experts' beliefs. For application-driven development and validation, we evaluate our approach in precision and robustness against a benchmark on synthetic datasets generated from the expert-knowledge simulation. Finally, we will demonstrate the learning results and findings with different encoding and levels of granularity on clinical datasets from the Swiss Paraplegic Center of secondary conditions in SCI individuals.

Poster presentation 34

Sex differences in intensive care mortality during infancy and childhood: A study of English linked administrative data

***Ofran Almossawi**, UCL/Great Ormond; Katie Harron, UCL; Richard Feltbower, University of Leeds, Simon Nadel, Imperial; Bianca De Stavoal, UCL*

Introduction: Infant and child mortality are important indicators of population health. Despite their sizeable reductions over the last two decades, distinct inequalities remain, as child mortality rates vary not only by age and income status, but also by sex, with higher mortality rates generally found in males.

In the population of children admitted to Paediatric Intensive Care Units (PICU), our previous research on infants has shown that mortality in PICU is higher for females relative to males, despite more male infants being admitted to PICU.

Aim: To examine empirically the factors that may play an important role in PICU admission and mortality of children in PICU.

Methods: Using data from Paediatric Intensive Care Audit Network (PICANet) and Hospital Admission Episodes, we created a linked cohort of 70,000 children born in England between 2002 and 2019, who had at least one PICU admission between January 2010 and December 2019. The linked cohort contains maternal, birth, and children's hospital admission records (prior to PICU admission), and mortality records from the Office of National Statistics.

Restricting the focus on the PICU population creates associations between predictors of admission (including sex), an example of collider bias. We used directed acyclic graphs to specify our assumptions and identify the non-causal paths linking sex and mortality in PICU. To overcome the high dimensionality of the identified potential confounders, we used machine learning methods to specify the propensity score function used to estimate the average causal effect of being female in the females population in PICU.

Poster presentation 35

Causal machine learning for heterogeneous treatment effects in the presence of missing outcome data

***Matt Pryce**, London School of Hygiene and Tropical Medicine*

In recent years, with researchers keen to fully utilise large observational datasets, the interest in exploring complex causal relationships has grown. However, to estimate such relationships, estimation techniques need to be robust enough to capture this complexity but also capable of handling the practical issues presented by observational data. To do this, the field of causal machine learning has proposed various ways in which machine learning, originally designed for predictive purposes can be used within a causal framework to estimate heterogenous treatment effects accurately and efficiently. One such method is the DR-learner, which estimates the conditional average treatment effect (CATE) by leaning on orthogonal learning principles, utilising a combination of propensity score re-weighting and direct regression estimation. This approach has been shown to be highly effective when used on complete data. However, in the presence of missing outcome data, the method can struggle to meet the required missingness assumptions, and even if the assumptions can be met, the missingness leads to slow model convergence. To address this, we propose an extension of the DR-learner, incorporating censoring weights while still meeting orthogonal learning principles. Our robust solution for estimating heterogenous treatment effects in the presence of missing outcome data holds improved convergence properties as well as additional practical benefits when missingness is present, both of which are demonstrated through the use of convergence bounds and simulation studies.

Poster presentation 36

Identifiability and estimation of exponential family models under missing not at random mechanism

***Anna Guo**, Emory University; Razieh Nabi, Emory University; Jiwei Zhao, University of Wisconsion-Madison*

Conducting valid statistical inference is challenging in the presence of nonignorable missing mechanisms, often referred to as missing-not-at-random or MNAR for short. In this work, we consider a MNAR missingness mechanism, most appropriate in cross-sectional studies. This model is a supermodel of several popular models, including permutation model (Robins 1997), block-conditional MAR model (Zhou et al. 2010), and block-parallel model (Mohan et al. 2013), thus making it less stringent in terms of underlying statistical assumptions. The underlying complete-data distribution in our cross-sectional MNAR model is not nonparametrically identifiable from the partially observed data. We establish sufficient conditions for identification of the target law by restricting our attention to the exponential family distributions. Unlike most prior work, all variables in our model can be subject to missingness, i.e., our results do not rely on the presence of fully observed variables. Borrowing the graphical model toolkit, we propose methods for testing the independence restrictions encoded in our model. If the test result suggests further independence restrictions in the model, we show that the model is nonparametrically identifiable. We also adopt a conditional likelihood approach for independence tests via estimating pairwise odds ratios. Further, we suggest the generalized method of moments for target law estimation. Statistical properties of the estimators are established.

Poster presentation 37

Counterfactual inference in post-treatment subpopulations: A Bayesian nonparametric approach

Johan de Aguas, *Department of Child Health and Development at the Norwegian Institute of Public Health; Department of Mathematics at University of Oslo; Guido Biele, Department of Child Health and Development at the Norwegian Institute of Public Health; Johan Pensar, Department of Mathematics at University of Oslo*

This paper investigates the problem of estimating counterfactual effects in subpopulations defined by conditional statements containing lagged or post-treatment variables. In general, such unit-specific counterfactual effects might not be identifiable with observational nor experimental data, as they involve alternative results under hypothetical scenarios that might conflict with real-world evidence. We review the general structural conditions for nonparametric identifiability of unit-specific counterfactual distributions using the twin network formulation. We further provide a two-step nonparametric estimating procedure that leverages soft-Bayesian additive regression trees (soft-BART) to perform abduction of relevant background noise variables, prediction of the outcome, and approximate computation of counterfactual contrasts. In a simulation study, our method reveals state-of-the-art performance, matching related precision benchmarks for heterogeneous treatment effects. We apply it to the problem of estimating the effect of medical treatment for attention deficit hyperactivity disorder (ADHD) on school performance in the subpopulation of Norwegian schoolchildren with comorbid disorders, including the case when comorbid diagnosis happened after the onset of treatment.

Poster presentation 38

User's guide to causal discovery on healthcare in the presence of hidden confounders

***Mirthe M. van Diepen**, Institute for Computing and Information Sciences; Ioan Gabriel Bucur, Institute for Computing and Information Sciences; Tom Claassen, Institute for Computing and Information Sciences*

Causal discovery can be beneficial to gain more knowledge about the true causal mechanism behind a model that is captured only in observed data, for instance, due to ethical reasons. For example, only observed data on high-risk thoracic aortic surgery was collected at the St. Antonius hospital in Nieuwegein, The Netherlands. Getting a better understanding of the risk factors for adverse outcomes of this surgery would be extremely helpful to optimize patient selection such that the mortality is minimized and fewer patients suffer from delirium or other high-risk complications after the surgery. The main challenges of this data set are (1) small sample size, (2) missing values, (3) context variables and time-dependent variables (different phases in the perioperative period), (4) unknown causal structure (there might be cycles or confounders), and (5) a complex combination of different variables, both discrete and continuous. We will suggest how to face these challenges, for example, what to consider when dealing with missing values or when choosing a causal discovery method. We will discuss the impact of the choices for the hyperparameters of the causal discovery methods, how we can use the hyperparameters to make the outcome more robust for small sample sizes and use the hyperparameters to deal with time-dependent variables.

Poster presentation 39

Association of mRNA COVID-19 vaccination during pregnancy with placental weight and birthweight: A population-based registry-linkage study in Norway

***Anteneh Desalegn**, PharmacoEpidemiology and Drug Safety Research Group, Department of Pharmacy, UiO; Angela Lupattelli, PharmacoEpidemiology and Drug Safety Research Group, Department of Pharmacy, UiO.*

Background: COVID-19 vaccination appears to be not only safe but also beneficial to both the pregnant mother and their offspring. However, there is scarcity of safety data especially during the first trimester. Objective: To determine the association between complete mRNA COVID-19 vaccination and the risk of low placental weight, birthweight to placental weight ratio, and small for gestational age according to timing of vaccination in pregnancy. Method: We conducted target trial emulation of a population-based registry-linkage in Norway between January 1st and December 31st, 2021. Time zero was specified based on time (week) of first vaccination from the last menstrual period (LMP), and corresponding controls are chosen weekly from LMP. Inverse probability time weighting is used to achieve the assumption of conditional exchangeability at baseline. The Marginal causal effects are estimated as weighted averages of conditional effects using Cox proportional-hazards model. Result: Among 50,243 eligible births, 6,711(13%) were from pregnant mothers vaccinated for at least one dose while 4,886 (10%) were from completely vaccinated for mRNA COVID-19 vaccine. About 6,098 (12%), 4,317 (9%), and 3,627 (7%) of the births had low placenta-weight, high birthweight-placenta weight ratio, and small for gestational age, respectively. Most (92.6%) were vaccinated either during the 2nd or 3rd trimester of pregnancy. Vaccination during pregnancy was associated differently depending on the timing of vaccination: positive association trend for early to mid-pregnancy while none or negative for latter stages. Conclusion: The finding suggests that the risk of m-RNA COVID-19 vaccination during pregnancy depends on the timing of vaccination.

Poster presentation 40

What makes forest-based heterogeneous treatment effect estimators work?

Susanne Dandl, *LMU Munich; Andreas Bender, LMU Munich; Christian Haslinger, University of Zurich; Torsten Hothorn, University of Zurich; Heidi Seibold, IGDORE; Erik Sverdrup, Stanford University; Stefan Wager, Stanford University; Achim Zeileis, University of Innsbruck*

Postpartum hemorrhage (PPH), defined as blood loss >= 500 ml within 24 hours of delivery by the WHO, is a short-term complication associated with maternal morbidity and mortality. To reduce the risk of PPH, it is vital to assess for each woman individually how the mode of delivery (cesarean section vs. vaginal delivery) affects postpartum blood loss (PBL). Estimating heterogeneous treatment effects (HTE) has two major challenges: First, PBL is right-skewed and interval-censored. Second, only observational data are available because the mode of delivery cannot be randomized. Since potential risk factors such as maternal age or estimated birth weight influence both the mode of delivery and PBL, confounding effects are expected. Model-based forests (MOBs) (Seibold, Zeileis, and Hothorn (2018)) and causal forests (CFs) (Athey, Tibshirani, and Wager (2019)) hold potential for this situation: MOBs combine the parametric modeling framework with random forests to estimate HTEs for a variety of outcome types, whereas CFs employ Robinson's (1988) orthogonalization strategy to increase robustness to confounding. Neither method is directly applicable to estimating HTEs for PLB: MOBs are not robust to confounding, and CFs are not applicable to interval-censored and right-skewed outcomes. To overcome these limitations, we combine the best of both worlds and identify how MOB's flexibility can be paired with CF's orthogonalization. We show that our blended approaches reduce confounding effects in a simulation study for various outcome distributions. To demonstrate this in practice, we employ our approach for estimating HTEs of cesarean section on PBL using observational data.

Poster presentation 41

Evaluating the impact of Fablabs on creativity and grit of high school students: An RCT with noncompliance

***Veronica Ballerini**, University of Florence; Martina Francesca Ferracane, European University Institute; Fiammetta Menchetti, University of Florence; Silvia Noirjean, University of Florence*

The technological revolution happened during the last decades has been deeply changing the labour market and, in turn, the technical, cognitive, and non-cognitive skills required of the workers. However, the educational systems are struggling to adapt to this new reality. In this study, we conduct an RCT randomly exposing students from five different Italian high schools to "digital fabrication" activities taught by FabLabs, to assess their impact on creativity and grit. Assignment is randomized at a class level; only students of classes assigned to the treatment are given the opportunity to participate in the Fablabs' activities. We address noncompliance adopting a Principal Stratification strategy; in this case, two latent subpopulations, i.e., the "Compliers" and the "Never-Takers", are defined by the potential compliance behaviour under treatment. The principal stratum membership is informed by several pre-treatment variables; individual characteristics, variables related to the family background, and the value of the outcome variable itself, measured at the baseline. We adopt a Bayesian approach to inference. Preliminary results on self-assessed creativity and grit show positive significant effects of the implemented program for Compliers; in the case of grit, results are robust to different models' specifications. Further analyses are needed and currently in progress; among others, the assessment of the impact on creativity using data collected administering survey questions from the "PISA creative thinking framework" (OECD, 2021) and from the Alternative Uses Test (Guilford, 1967).

Poster presentation 42

Causal inference in dementia prevention using data from the Canadian longitudinal study on aging

***Vicky Chang**, Western University; Abolfazl Avan, Mark Daley, Vladimir Hachinski; Abolfazl Avan; Mark Daley; Vladimir Hachinski*

With an aging global population, dementia incidence is growing at an alarming rate. Current data shows that at least 40% of dementias are attributable to modifiable risk factors and thus preventable. To inform policymakers and clinicians which risk factors should be targeted to have the greatest impact on dementia prevention, we require a causal analysis of the data. Our research aims to apply the principles of causality inference to determine the relative weighting of protective and harmful factors in the development of cognitive impairment. We are using data from the Canadian Longitudinal Study on Aging (CLSA), which follows 50, 000 individuals between ages 45 and 85 over 20 years. Demographic, socioeconomic, lifestyle, and clinical data were acquired through interviews, and cognitive function was assessed at baseline and follow-ups. As part of the Dementia Prevention Initiative at Western University, we are working with a multidisciplinary team of investigators, and integrating their expertise to inform how the various risk factors are connected in our causal diagram. Our approach uses the CLSA data and the DoWhy Python library to create a causal model from the data. With our causal analysis, we aim to answer questions like "Which risk factor is the root cause of Person A developing cognitive impairment, and why did Person B experience a different result?" We will run sensitivity and robustness checks on the estimates obtained from several different models accounting for various combinations of risk factors to optimize our predictive accuracy.

Poster presentation 43

Convolutional neural networks for valid and efficient causal inference

Mohammad Ghasempour, *Umea university; Niloofar Moosavi, Umea university; Xavier de Luna, Umea university*

Convolutional neural networks (CNN) have been successful in machine learning applications including image classification. When it comes to images, their success relies on their ability to consider the space invariant local features in the data. Here, we consider the use of CNN to fit nuisance models in semiparametric estimation of a one dimensional causal parameter: the average causal effect of a binary treatment. In this setting, nuisance models are functions of pre-treatment covariates that need to be controlled for. In an application where we want to estimate the effect of early retirement on a health outcome, we propose to use CNN to control for time-structured covariates. Thus, CNN is used when fitting nuisance models explaining the treatment assignment and the outcome. These fits are then combined into an augmented inverse probability weighting estimator yielding efficient and uniformly valid inference. Theoretically, we contribute by providing rates of convergence for CNN equipped with the rectified linear unit activation function and compare it to an existing result for feedforward neural networks. We also show when those rates guarantee uniformly valid inference for the proposed estimator. A Monte Carlo study is provided where the performance of the proposed estimator is evaluated and compared with other strategies. Finally, we give results on a study of the effect of early retirement on later hospitalization using a database covering the whole Swedish population.

Poster presentation 44

Continuous-time mediation analysis for repeated mediators and outcomes

Kateline Le Bourdonnec, *Université Bordeaux; Cécilia Samieri, Université de Bordeaux ; Linda Valeri, Columbia Univ. New-York ; Cécile Proust-Lima, Université de Bordeaux*

Mediation analysis focuses on the underlying causal mechanisms between an exposure, an outcome, and an intermediate variable called mediator. Initially developed for cross-sectional studies, it has been extended to the framework of longitudinal data by discretizing the assessment times of mediator/outcome. Yet, processes in play in longitudinal studies are usually defined in continuous-time and measured at irregular and subject-specific visits. This is the case in dementia research when interested in causal mechanisms involving neurodegeneration and cognitive functioning. Our objective was to propose a methodology to estimate the causal mechanisms between a time-fixed exposure (X), a mediator process (Mt) and an outcome process (Yt) both measured repeatedly overtime in presence of a time-dependent confounding process (Lt).We considered two causal estimands, a path-specific effect and a stochastic intervention analogue to natural effect, and defined the identifiability assumptions required to get the estimands estimable. Then, we used a dynamic multivariate model based on differential equations for their estimation. We applied our method in a population-based cohort of cerebral aging to investigate the causal mechanism between a genetic factor (APOE4) and cognitive functioning potentially mediated by neurodegeneration and confounded by vascular brain lesions.

Poster presentation 45

Sensitivity Analysis with the R^2-calculus

Tobias Freidling, *University of Cambridge; Qingyuan Zhao, University of Cambridge*

Causal inference necessarily relies upon untestable assumptions, such as the absence of unmeasured confounders; hence, it is crucial to assess the robustness of obtained results to violations of identification assumptions. However, such sensitivity analysis is only occasionally undertaken in practice as many existing methods only apply to relatively simple models and their results are often difficult to interpret. We take a more flexible approach to sensitivity analysis and view it as a constrained stochastic optimization problem. This work focuses on linear models with an unmeasured confounder and a potential instrument. In this setting, the R^2-calculus – a set of algebraic rules that relates different (partial) R^2-values and correlations – emerges as the key tool for sensitivity analysis. It can be applied to identify the bias of the family of k-class estimators, which includes the popular OLS and TSLS estimators, as well as construct sensitivity models flexibly. For instance, practitioners can specify their assumptions on the unmeasured confounder by comparing its influence on treatment/outcome with an observed variable. We further address the problem of constructing sensitivity intervals which cover the range of parameters under the partially identified sensitivity model with high probability and generalize the concept of confidence intervals. Since the heuristic "plug-in" sensitivity interval may not have any confidence guarantees, we instead propose a bootstrap approach which performs well in numerical simulations. We illustrate the proposed methods with a real study on the causal effect of education on earnings and provide user-friendly visualization tools.

Poster presentation 46

A semi-parametric model for conditional local independence testing applied to causal graph learning

***Myrto Limnios**, Department of Mathematical Sciences, University of Copenhagen; Niels R. Hansen, Department of Mathematical Sciences, University of Copenhagen*

We are interested in estimating the causal effect of a time-continuous treatment on a set of diseases among a studied population. Indeed, data collection techniques store massive high-dimensional time-dependent structures in many applications requiring a causal understanding.

By considering these markers as time-continuous event processes, existing learning algorithms are applicable but rely on a model for the event processes or the test statistic, e.g. algorithms for directed mixed graphs [Mogensen et al. (2018,2020)], and the Causal Analysis algorithm [Meek (2014)].

In this work, we propose a semi-parametric model to learn the structure of the underlying causal graph. We approximate both the unknown intensity process and a particular version of the conditional local independence test from [Christgau et al. (2022)] using their respective Volterra expansion. The latter is defined as a linear expansion of tensor products of kernels functions composed of stochastic integrals w.r.t. the observed event processes. A lasso optimization formulation estimates the involved parameters, and under some assumptions, nonasymptotic probabilistic bounds are investigated to obtain data-driven optimal weights for the penalty terms, following e.g. [Bacry et al. (2020), Hansen et al. (2015)]. These results also guarantee the quality of the estimated processes using the Volterra expansion model w.r.t. minimizing the empirical squared-risk function.

Once the optimal parameters are estimated for a particular node of the graph, a practical implementation of the statistical test based on cross-validation and sample splitting is proposed adapted from [Christgau et al. (2022)], to be plugged-in the aforementioned algorithms.

Poster presentation 47

Fuzzy difference-in-differences for spatiotemporal data

***Andrej Srakar**, Institute for Economic Research (IER), Ljubljana and University of Ljubljana; Marilena Vecco, Burgundy School of Business, Dijon, France*

Difference-in-differences (DiD) literature is a fast growing field in econometrics. Topics such as presence of serial correlation, clustered standard errors, arbitrary covariance structures, parallel growth assumption, negative controls, matching, synthetic control, semi- and nonparametrics, nonlinear models, multiple and continuous treatments and staggered treatment adoption have been subject to recent research. In previous contributions DiD has been extended to spatial data (Delgado and Florax, 2015). We extend this in a treatment effect with network interference context by controlling for violated stable unit treatment values assumption (SUTVA), inherent for such analysis but seldom controlled so far and in a spatiotemporal autoregressive setting. We develop a time-corrected Wald ratio DiD estimator combining fuzzy DiD approach with extensions to changes-in-changes estimation (Athey and Imbens, 2006; De Chaisemartin and D'Haultfoueille, 2017) with a graphon random graph specification (Li and Wager, 2022). This allows asymptotic analysis in terms of bounds on the moments and Stein approaches including Monte Carlo simulation results. In an application, we study causal effects of the yearly Venice carnival, being able to isolate the effect respective to other competing large events in Venice in the studied period which solves one of main open problems in economic impact analysis. In conclusion, we consider extensions using spillover double robust DiD (Clarke, 2017) and Bayesian approaches.

Poster presentation 48

The causal effect of retirement and mental health: Partial identification with panel

***Eduardo Fe**, University of Manchester*

We study how retirement affects mental health using a partial identification framework for panel data. We consider identifying assumptions that restrict the process of selection, bind the amount of variation in counterfactual moments or limit the amount of interference within units along the time dimension. The identification regions that we estimate are then transformed into quantitative policy recommendations by adopting a minimax-regret decision rule. We find that retirement can have a moderate, positive effect on younger individuals whose pre-retirement mental health is below average of the corresponding age cohort. Our minimax-regret analysis further suggests that while retirement can be delayed without affecting the mental health of the majority of the population, it is advisable to facilitate the retirement of people with poor mental health.

Poster presentation 49

Evaluating changes in mental well-being after the Universal credit welfare reform: A Bayesian spatial interrupted time series model for England

***Connor Gascoigne**, Imperial College London; Marta Blangiardo, Imperial College London; James Kirkbride, University College London; Sara Geneletti, London School of Economics and Political Science; Annie Jeffery, University College London; Zejing Shao, University College London; Gianluca Baio, University College London*

In 2013 the welfare reform, Universal Credit (UC), was introduced to the UK with the purpose of replacing six existing benefits. The aim of this study is to evaluate if there are any changes in psychological distress after a contextual awareness of UC in England over a study period of 2009 to 2021 as this spans the UC rollout which started in 2013 and was present in all English local authorities by 2016. Within the UK's household longitudinal panel survey, we use the response to the General Health Questionnare-12 to define psychological distress. We define our control and exposed population using employment status and using the propensity score to account for several important socioeconomic confounders, as well as geographical location. We fit an interrupted time series model via a Bayesian hierarchical framework that captures any spatial dependencies between geographical locations and at the same time is flexible enough to account for spatial departures from the general temporal trend. Our results suggest a strong difference in the prevalence of psychological distress over the course of the study between the control (19.74% average with a 95CI of [16.94%, 22.84%]) and exposed populations (31.42% with a 95CI of [28.64%, 34.26%]). We see an increase in the prevalence of psychological distress in the year immediately before a contextual awareness to UC, followed by a decrease in the year immediately after. When looking at different administration levels and within different demographics (deprivation index and ethnicity), we see the same pattern.