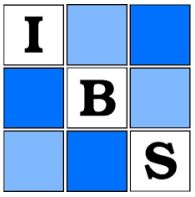

Book of Abstracts

Nordic-Baltic Biometric Conference 2025

June 9-12, Oslo, Norway

Keynote lectures

Domain adaptation techniques for intensive care databases

***Peter Bühlmann**, ETH Zurich*

We discuss the challenge of predicting individual patient outcomes in intensive care units (ICUs). While several medical databases offer vast amounts of data, integrating this information for specific ICUs or patients remains complex. We explore conceptual modeling paradigms that facilitate generalization and domain adaptation to new units, leveraging Empirical Bayes methods and Causal-Inspired Machine Learning. Through empirical validation, we uncover valuable insights into the effectiveness of these approaches.

SJS lecture: Derivation of and inference about partial identification bounds in small samples

***Erin Gabriel**, University of Copenhagen*

There has been renewed interest in the use and derivation of partial identification bounds. One way of deriving such bounds is linear programming. This may result in symbolic or numeric partial identification bounds. The latter are dependent on data and must be recalculated for each new dataset. Particularly in the small sample setting, this can cause a problem, as the derivation and estimation of bounds are not separate. In this talk, I will highlight the derivation of symbolic bounds using linear programming and their pros/cons in contrast to numeric partial identification bounds. I will then outline a novel but straightforward method for accounting for sampling variability that does not require assumptions about the true value being away from the boundary. I will demonstrate the method and compare it to the bootstrap and the m-out-of-n bootstrap, which have both been suggested previously in the literature. Finally, I will outline some next directions that could improve estimation and inference, by restricting the parameter space by the known observable constraints from the graph used to derive the bounds.

Some perspectives on models and their parametrisations

***Heather Battey**, Imperial College London*

Two broad positions within statistics define a treatment effect, on the one hand, as a parameter of a statistical model, and on the other, as an appropriate population-level difference in outcomes or counterfactual outcomes under the different treatment regimes. I will start by presenting some simple but consequential perspectives on the two formulations, contrasting the answers under a fictitious idealisation that isolates key issues in an illuminating form.
The most common objection to model-based formulations is the possibility of misspecification, and cautious agnosticism often entails the introduction of a large number of parameters. These may be considered in different ways, depending on context, e.g. as fixed arbitrary constants, as iid random variables, or as determined by potentially relevant intrinsic variables. I will discuss some of the implications of each formulation with reference to work with Nancy Reid on the role of parametrisation in models with a misspecified nuisance component, and with David Cox on confidence sets of models.

Invited session: Statistics and AI

Differential meometry in Machine learning: from generative modeling to Bayesian inference and beyond...

***Georgios Arvanitidis**, Technical University of Denmark (DTU)*

A common assumption in machine learning is that data lie near a low-dimensional manifold embedded in a higher-dimensional space, though this manifold is typically unknown. In this talk, we will present techniques based on deep generative models that allow us to learn the geometric structure of the data manifold. This, in turn, enables applications ranging from statistical modeling on nonlinear geometries to the design of robot controllers with stability guarantees. Beyond this classical use case of manifolds in machine learning, we will also explore how considering the geometry of the loss landscape in deep neural networks can enhance approximate Bayesian inference methods.

Training-free guidance of generative AI models as Bayesian inference

***Fredrik Lindsten**, Linköping University*

Diffusion and flow-based models have emerged as a powerful tool in generative AI, enabling the creation of high-quality images and audio, but also other forms of data appearing in the natural sciences, such as molecules and materials. These models can be "guided" to simulate from desired conditional distributions, for instance to condition on a desired property of generated molecules in the context of drug discovery. However, standard guidance techniques require paired data, expensive training or fine-tuning, and is limited to specific types of conditioning. Recent advancements have introduced training-free guidance techniques that allow for more flexible and efficient control over the generation process. This is based on interpreting the conditional generation as a Bayesian inference problem, where the generative model acts as a prior and the conditioning is incorporated via a likelihood term. This talk will explore the fundamentals of diffusion and flow-based generative models and introduce some of our recently developed Bayesian inference methods for training-free guidance of such models.

Is it enough to include the output after training a supervised model?

***Oskar Allerbo**, Royal Institute of Technology (KTH)*

Given the success of unsupervised learning, can supervised models also be trained without using the information in the output? In this talk, we demonstrate that this is indeed possible. The key step is to formulate the model as a smoother, i.e. with predictions on the form f=Sy, and to construct the smoother matrix, S, independently of the output, y. We present a simple model selection criterion based on the distribution of the out-of-sample predictions and show that, in contrast to cross-validation, this criterion can be used also without access to y. We demonstrate on real and synthetic data that y-free versions of linear and kernel ridge regression, smoothing splines, and neural networks perform similarly to their standard, y-based, versions and, most importantly, significantly better than random guessing.

Invited session: Functional data analysis

Analysis of ice coverage in Baffin Bay from 1979 to 2023 as a functional time series

***Helle Sørensen**, University of Copenhagen; Susanne Ditlevsen, University of Copenhagen; Mads Peter Heide Jørgensen, Greenland Institute of Natural Resources*

It is well known that the waters around Greenland are changing rapidly, with the reduction and thinning of sea ice cover being one of the most pronounced changes, and with severe impacts on the marine ecosystem. We study ice cover in Baffin Bay between northern Canada and Greenland, a basin which is typically covered with ice in winter and ice-free in summer. Data consist of daily ice cover (in square meters) in the period from 1979 to 2023. We consider the data as a functional time series with each year providing a curve used as the data unit. With methods from functional data analysis, we examine if there is a shift around year 2000 (as detected in other regions) and whether the ecosystem resets itself every year, corresponding to independence over the years.

Double robust estimation of functional outcomes with data missing at random

***Felix Ecker**, Swedish University of Agricultural Science*

This talk presents semi-parametric estimators for the mean of functional outcomes in situations where some of these outcomes are missing at random. That is, the missingness mechanism depends only on some fully observed covariates. We present two estimators for the functional mean, using working models for the functional outcome given the covariates, and the probability of missingness given the covariates. We establish that both estimators have Gaussian processes as their limiting distributions and explicitly give their covariance functions. One of the estimators is double robust in the sense that the limiting distribution holds whenever at least one of the nuisance models is correctly specified. These results allow us to present simultaneous confidence bands for the mean function with asymptotically guaranteed coverage. A Monte Carlo study shows the finite sample properties of the proposed functional estimators and their associated simultaneous inference. The talk further discusses how these estimators can be used to target counterfactual functional outcomes in the context of causal inference.

Functional response regression models with latent variables

***Valeria Vitelli**, University of Oslo (UiO)*

In functional regression problems, the response curves are often observed partially, and incompleteness of observation might refer to the fact that the curves are observed sparsely, the most popular case in the literature, or completely latent. We focus on the latter case, and specifically on situations in which the values of the response curves at each time point are not observed directly, but via their discretized versions in the domain, either sparsely or densely. Among the many potential applications of this general framework, in a clinical setting one could think of the intensity of a condition being measured via the absence/presence of symptoms along time. This class of problems can be handled via the family of generalized functional regression models, which however rely on a viable approach to the estimation of the functional covariance operator. We propose a novel Function-on-Scalar Regression (FoSR) model setting, where the latent response variable is a latent Gaussian random element taking values in a separable Hilbert space, and we only observe its realization as a sequence of correlated binary observations. The smoothness constraints are put on the latent response curves instead of constraining the regression coefficients and eigen-functions directly, which allows easily controlling the smoothness level of functional unknowns. We also propose a practical computational framework for maximum likelihood inference via the parameter expansion technique, which flexibly handles both non-equally spaced and missing observations effectively, and is based on an Adaptive Monte Carlo Expectation-Maximization (AMCEM) algorithm. We showcase the method performance on multiple simulated scenarios, and on a case study concerning psychiatric symptoms.

Invited session: Causal inference

Inference on variable importance measures for heterogeneous treatment effects

***Paweł Morzywolek**, University of Washington, USA*

Recent years have seen a growing interest in quantifying treatment effect heterogeneity, which is vital for supporting individualized decision-making. Though black-box machine learning approaches might optimally predict treatment effect heterogeneity, in high-risk domains such as medicine, decision makers often hesitate to rely on decision support systems without understanding the underlying rationale behind the recommendations. Hence, it is crucial to offer insights into which variables best predict individualized treatment effects. Motivated by these considerations, we present model agnostic variable importance measures for heterogeneous treatment effects. Our approach builds on recent developments in semiparametric theory for pathwise differentiable function-valued parameters and is valid even when flexible black-box algorithms are employed to quantify treatment effect heterogeneity.

Causal inference with continuous exposures

***Michael Schomaker**, Ludwig Maximilian University of Munich, Germany*

Often, an exposure (treatment, intervention) of interest is continuous and measured over multiple time points. A typical estimand in this case is the longitudinal causal dose-response curve (CDRC). For example, in pharmacoepidemiology, one may be interested in how outcomes of people living with -and treated for- HIV, such as viral failure, would vary for different time-varying exposures such as different antiretroviral drug concentration trajectories (ultimately to determine the therapeutic window for maintaining drug concentrations within an effective and safe range).
A challenge for doing causal inference with continuous exposures is that the so-called positivity assumption is typically violated. The assumption requires positive conditional exposure densities at all time points, at all exposure trajectories of interest. We present 3 possible strategies to address this: 1) using standard g-computation approaches that are used for binary exposures, without any modifications, relying purely on extrapolations; 2) developing projection functions, which reweigh and redefine the CDRC based on functions of the conditional support for the respective exposure strategy: with these functions, we obtain the desired dose-response curve in areas of enough support, and otherwise another estimand that does not require the positivity assumption; 3) an individual, data-adaptive strategy that sticks to the exposure trajectory of interest as long as possible, and uses the closest "most-feasible" exposure value otherwise. We develop g-computation type plug-in estimators for strategies 2 and 3.
Simulations show in which situations a standard g-computation approach (strategy 1) is appropriate, and in which it leads to bias and how then strategy 2 recovers the alternative estimand of interest, and strategy 3 reduces bias (while maintaining interpretability).
All ideas are illustrated with longitudinal data from HIV positive children treated with an efavirenz-based regimen as part of the CHAPAS-3 trial, which enrolled children <13 years in Zambia/Uganda.

Formulating and comparing adherence strategies for sustained treatment: a nationwide case study of five years of endocrine therapy in patients with breast cancer

***Elise Dumas**, Federal Polytechnic School of Lausanne (EPFL), Switzerland*

Lack of adherence can reduce the effectiveness of beneficial treatments. However, the extent to which adherence affects outcomes is unclear in many settings. For example, is it enough to adhere to treatment for 80% or 90% of the prescribed days? Does it matter whether adherence is higher earlier or later in the treatment schedule? And is it particularly important not to miss consecutive days of treatment? In this work, we explore a methodology to emulate and compare different adherence strategies. Specifically, we use a framework that incorporates explicit grace periods and regimes that depend on natural treatment patterns. Our work is motivated by a clinical question concerning women with early-stage hormone receptor-positive

breast cancer, for whom daily endocrine therapy is prescribed for five to ten years. In these patients, young age is associated with both an increased risk of cancer recurrence and suboptimal adherence to endocrine therapy. Using French nationwide claims data, we applied the proposed methods to compare the survival benefits achievable in patients with breast cancer by implementing different adherence strategies to endocrine therapy for each age group. We emulated three different adherence strategies that allowed for gaps in treatment of no more than one, three, or six consecutive months. A total of 121 601 patients were included in the analyses. In patients aged 34 years or younger, strict ET adherence ($\leq$1-month gaps) improved 5-year DFS by 4.3 percentage-points, (95% confidence interval (CI): 2.6-7.2) compared to observed adherence. In this age group, ET adherence strategies allowing for $\leq$3-month and $\leq$6-month gaps reduced the 5-year DFS benefit to 1.3 (95% CI: 0.2-3.7) and 1.0 (95% CI: -0.2-3.4) percentage-points, respectively. In contrast, DFS benefits of improved ET persistence in patients after 50 years old did not exceed 1.8 percentage-points, compared to observed persistence, regardless of the length of gaps allowed. Our results show that young women would benefit substantially from stricter adherence to endocrine therapy, with treatment breaks never exceeding one month, highlighting the need for tailored strategies to improve treatment adherence in this population.

Invited session: Statistics for personalized medicine

New tools for association testing, estimating heritability and constructing polygenic scores

***Doug Speed**, Aarhus University, Denmark*

I will describe three new GWAS tools. The first is LDAK-KVIK, a tool for mixed-model association testing of quantitative and binary traits. LDAK-KVIK is computationally efficient, requiring less than 6 CPU hours and 8Gb memory to analyse genome-wide data for 400k individuals. Further, it is powerful, finding more associations than the existing tools BOLT-LMM, REGENIE and fastGWA. The second is TetraHer, a tool for estimating the heritability of diseases, that allows for covariates and ascertainment. Applied to data from UK Biobank, TetraHer identifies 107 ICD-10 diseases with significant heritability. The third is a revised version of MegaPRS, an existing tool for constructing polygenic scores. Compared to original MegaPRS, the new version both improves accuracy, and is able to integrate data from multiple traits and populations.

Bayesian variable selection for multi-omics modelling and outcome prediction in precision oncology

***Manuela Zucknick**, University of Oslo, Norway*

Large-scale cancer pharmacogenomic screening experiments profile cancer cell lines or patient-derived cells versus hundreds of individual drug compounds or drug combinations. The aim of these in vitro studies is to use the genomic profiles of the cell samples together with information about the drugs to predict the response to a particular treatment. Unfortunately, the in vitro cell viability experiments that measure drug response are technically challenging and expensive, and even the largest drug screens only include a few dozen or at most hundreds of cell samples. This means that we are in a multi-response regression setting with few samples and high-dimensional heterogeneous (multi-omics) input data.
In this challenging setting we aim to enhance model performance by leveraging data structure, designing structured priors to combine data sources, borrow information across correlated response variables, and integrate external biological knowledge, such as drug target pathways. Our approach involves a multivariate Bayesian variable and covariance selection setup with several extensions. One such extension utilizes a Markov random field prior for latent variable selection to exploit known structures, like molecular pathways linked to specific drugs. Another recent development incorporates interaction effects with modifying variables to account for heterogeneity in the effects of individual omics input variables on drug response outcomes.

Genetic Insights into Disease Onset and Progression

***Zhiyu Yang**, University of Helsinki, Finland*

Understanding disease progression is of significant biological and clinical interest. While the genetic underpinnings of disease susceptibility have been extensively studied, much less is known about the genetics of disease progression and its relationship to susceptibility. To address this gap, we defined disease-specific mortality as a proxy for progression in ten common diseases and systematically compared the genetic architectures of susceptibility and progression across seven major biobanks. Our analyses revealed limited overlap in genetic effects between disease susceptibility and disease-specific mortality.
We also identified several key challenges in studying the genetics of disease progression, particularly regarding the definition and measurement of progression phenotypes. To refine this, we leveraged trajectories of repeated clinical laboratory measurements in FinnGen to model the process of disease onset. Using linear mixed models and functional principal component analysis (fPCA), we investigated the genetic determinants of trajectory shapes for disease-relevant lab values and compared these to the genetics of individual mean lab value levels.

Our findings indicate substantial overlap in genetic influences on both trajectory shape and mean levels of disease-relevant lab value, suggesting potential horizontal pleiotropy. However, we also identified trajectory-specific genetic variants, providing novel insights into the biology of disease development and offering potential for improved disease prediction.

Invited session: Statistics in ecology

Fish stock assessment

***Olav Nikolai Breivik**, Norwegian Computing Center, Norway*

In this talk, I will present a state space fish stock assessment model and highlight recent research aimed at improving assessment quality. The model is currently the most widely used for setting quota advice in Europe. I will focus on how data from commercial catches and scientific surveys are used to inform the model, and elaborate on how spatio-temporal structures in survey data can be propagated into the assessment.

Latent Markov models in ecology: a unifying perspective

***Jan-Ole Koslik**, Bielefeld University, Germany*

Latent (or hidden) Markov models are powerful tools for analysing time series or other sequential data that depend on underlying but unobserved states. Owing to their flexible hierarchical framework, which separates the noisy observation process from the latent Markovian state process, they have gained prominence across numerous empirical disciplines. In ecology in particular, they have become immensely popular, as ecological data collected over time are often characterised by indirect and incomplete observations linked to latent states.
For practitioners, however, the multitude of available inferential techniques and software packages can be overwhelming. To address this, I will introduce the LaMa R package, which adopts a unifying perspective across different combinations of discrete and continuous state and time formulations, while offering a modular, 'Lego-style' approach for crafting custom likelihood functions. Users can combine pre-built components or design new ones, enabling tailored solutions for complex ecological questions.
While LaMa is more generally applicable, this talk will focus on models with discrete (rather than continuous) state spaces, as hidden behavioural modes — such as resting, foraging, and travelling — are of particular interest in animal ecology. Specifically, I will present two ecological applications: modelling the movement of an African elephant using (discrete-time) hidden Markov models, and modelling surfacing times of minke whales using Markov-modulated Poisson processes.

Doing Large-scale Modeling of Species Distributions Properly

***Robert O'Hara**, Norwegian University of Science and Technology (NTNU), Norway*

Knowing the distributions of species is vital, for ecologists, conservation biologists and increasingly for local and national agencies. But the data on their distributions is a mess. I will describe our work to develop flexible models that can integrate different data types, and account for their inherent biases. This uses a point process within a state space framework. The actual distribution of individuals of a species is assumed to be points drawn from a random field, e.g. a log Gaussian Cox process. The observed data may be of different forms, binomial presence/absences, Poisson abundance counts, or "presence only" observations of locations where species were recorded. All of these have their own observation model, conditional on the random field. Models for data sets with different likelihoods can be combined into one big model, so that biases in some data sets can be corrected by comparison with unbiased (or less biased!) data.
I will describe how we have used this framework to model thousands of Norwegian species. This required the building of computational pipelines to fit semi-automated models to publicly available data. From this we can identify where there are hotspots of Norwegian biodiversity, and provide maps not just of distributions, but also of sampling effort, which will help guide improved monitoring of biodiversity.

Contributed session 1

Estimation of the effective reproduction number (R) for Covid-19 in Norway using splines and an SEIR model

***Magne Aldrin**, Norwegian Computing Center*

I present here a model for continuously estimating the daily effective reproduction number applied to the case of Covid-19 in Norway from the beginning of the epidemic in spring 2020 until February 2022. Important factors considered include the increasing level of official testing in the first months, then the introduction of self-testing and later the extensive use of vaccination. The full model includes a Susceptible-Exposed-Infected-Recovered (SEIR) model for the infection process, which is further linked to testing procedures and hospitalisation. The model uses smooth spline functions to estimate i) the daily development of the transmissibility parameter, ii) the proportion of infected individuals detected through testing, and iii) the proportion of Covid-19 cases requiring hospitalisation. Consequently, the model also estimates the expected prevalence and incidence of hospitalised Covid-19 patients. To estimate the model, publicly available daily data on the number of tests conducted, the proportion of positive tests, and the recorded incidence and prevalence of hospitalised patients are used.

Mapping hospital patient movements: implications for nosocomial infection transmission

***Birgitte Freiesleben de Blasio**, Norwegian Institute of Public Health, University of Oslo; Gianpaolo Scalia Tomba, Stockholm University, University of Oslo*

Understanding patient movement patterns is crucial for tracking the spread of hospital-acquired infections, with antimicrobial resistance (AMR) being a particular concern. In this study, we apply network analysis to a complete dataset of approximately 3 million patient registrations from South-Eastern Norway within a calendar year. We construct directed networks based on uninterrupted episodes of care, where nodes represent hospital wards and edges correspond to patient transfers. Descriptive analysis reveals significant temporal patterns in key statistics, including daily ward-to-ward movements and substantial heterogeneity in the distribution of movements per episode, patient, and ward. We classify wards based on their structural importance within the network using multiple centrality measures—degree, strength, closeness, betweenness, PageRank, and a novel measure, Thinning-Percolation. Thinning-Percolation focuses on the ability of each ward to be connected to broader parts of the network, respecting the direction of edges. Random thinning of edges can be interpreted as a transitioning from an accumulated network structure toward the instantaneous network on which dynamics occur. Our results identify a highly connected core of medical wards in large hospitals alongside surgical and specialised wards with regional responsibilities. We discuss our findings in the context of a large vancomycin-resistant enterococci (VRE) outbreak at Østfold Hospital, underscoring critical gaps in data availability. Finally, we highlight key considerations and limitations associated with using patient movement data for mathematical modelling of hospital infection dynamics.

A hierarchical bivariate logistic model for estimation of small-area prevalences

***Nicola Fitz-Simon**, University of Galway*

Data on health indicators collected via surveys with complex designs provide unbiased, low variance estimates of national-level proportions and ratios, using direct estimators that take sampling weights into account. To aid decision-making on where to target resources, data from such surveys are increasingly used to provide estimates at small-area level. When sample sizes are small direct estimators do not have good properties; small area estimation methods borrow strength across areas to provide more efficient estimates. Efficient estimation at small area level of ratios, such as what proportion of individuals with a disease have been previously diagnosed, may appear even more challenging due to smaller sample sizes. Previous applied studies have estimated small area subpopulation proportions by restricting the sample. However, outcomes

such as disease, diagnosis and treatment are often correlated at small area level. Thus, a bivariate logistic model including correlated random effects at area level can provide improved estimates of small area prevalences. By jointly modelling both the probability of disease and the probability of previous diagnosis, we demonstrate more efficient estimates of the probability of diagnosis given previous disease. Drawing on recent developments in small area estimation of proportions, we fit this model as the first stage of a two-stage estimation approach. We compare with the estimates from a model that does not leverage area-level correlations. We present the results of an application where the question of interest was to estimate at small area level the proportion of individuals with raised blood pressure who had been previously diagnosed.

Distributing disease burden from county to municipality level: developing a model for scaling global burden of disease results on YLL to local contexts

***Christian Madsen**, Norwegian Institute of Public Health; Carl Michael Baravelli; Liliana Vazquez Fernandez; Gunn Marit Aasvang; Ann Kristin Skrindo Knudsen; Anette Kocbach Bølling*

Background: Local health burden estimates are crucial for effective public health planning. The Global Burden of Disease (GBD) study provides national and, for some countries, subnational estimates of Years of Life Lost (YLL) estimates, but these may overlook local variations essential for targeted interventions. This study develops a method to downscale county-level GBD YLL estimates to the municipality level using statistical and machine learning approaches.
Methods: A cohort study using Norwegian population-based registries (2015-2019) applied models like XGBoost, Random Forest, Generalized Additive Models, and Bayesian regression to redistribute YLL rates. Initially, models were tested without additional covariates. The best-performing model, XGBoost, was then enhanced with demographic, socioeconomic, healthcare accessibility, environmental risk factors, health expenditures, and general practitioner consultation frequencies. Performance was evaluated using root mean squared error (RMSE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2).
Findings: XGBoost showed the highest predictive accuracy, with the lowest RMSE and highest stability. The full model, including contextual predictors, outperformed an age-and-sex-only model. Post-prediction adjustments using county-level scaling factors further improved accuracy.
Interpretation: This study presents a machine learning approach for downscaling YLL estimates to municipality levels, enhancing spatial granularity in disease burden assessments. This facilitates targeted public health interventions and efficient resource distribution. The framework offers a scalable solution for refining local health metrics, with potential applications beyond Norway. Further research is needed to validate the model in different settings and for different metrics, such as years lived with disability.

Characteristics and outcomes of hospitalised patients with cancer and COVID-19: a global prospective cohort study of 800,000 individuals from ISARIC

***Christiana Kartsonaki**, University of Oxford; Barbara Wanjiru Citarella, University of Oxford; Piero Olliaro, University of Oxford; Lance Turtle, University of Liverpool; Carlo Palmieri, University of Liverpool; Laura Merson, University of Oxford; ISARIC Clinical Characterisation Group*

Individuals with cancer are at a higher risk of adverse outcomes of SARS-CoV-2 infection, both because of the disease and of immunosuppression due to cancer treatments. However, risks may vary by cancer type, treatments received, age, and by country and time related to differences in cancer and COVID-19 treatment availability, immunity, and circulating variants. We analysed data on over 800,000 hospitalised patients with COVID-19 from 60 countries collected by the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC), with most patients from sites in South Africa and the United Kingdom. This is the largest global prospective cohort describing the characteristics and outcomes of patients with cancer and COVID-19 over the first two years of the pandemic. Cox regression was used to estimate hazard ratios of death associated with cancer. In total 36,212 patients had cancer reported, of whom 11,224 (31.0%) died during the COVID-19-associated hospitalisation. Cancer type and treatment information was available in a subset of patients. Metastatic cancer was reported for 867 patients and 481 were reported to have received chemotherapy. The most common cancer types were haematological (n=2102, of which 560 had leukaemia,

914 lymphoma, and 396 myeloma), breast (n=1493), prostate (n=1005), and lung cancer (n=820). Symptom prevalence on admission did not differ substantially between patients with and without cancer, and similar proportions met COVID-19 case definitions. Overall, cancer was associated with a higher risk of death during the COVID-19-associated hospitalisation (adjusted HR 1.25 [95% CI 1.22-1.28]) adjusted for age, age2, sex, and country. HRs were similar during the first (2020) and second year of the pandemic (2021). Relative risks of in-hospital death associated with cancer were higher for young patients, with the 20-30 age group having a HR of 3.87 (95% CI 2.72-5.49). Cancer is associated with a higher risk of death in hospitalised patients with COVID-19. Despite the availability of vaccines, patients with cancer remain at a high risk, likely due to less strong and less lasting immune responses, therefore booster vaccinations, therapeutics, and non-pharmacological interventions are likely to continue to be necessary. Our findings may inform clinical planning and patient management.

Contributed session 2

On estimating the Linfoot correlation

***Ulrich Halekoh**, University of Southern Denmark; Stéphanie van den Berg, University of Twente; Jacob vB. Hjelmborg, University of Southern Denmark; Andreas K. Jensen, University of Copenhagen; Sören Möller, University of Southern Denmark*

We propose to study dependence in structured data using the Linfoot correlation (Linfoot, 1957). This information theoretic based measure satisfies the seven properties stated by Renyi (1958). The most notable being independence if the measure is zero, transformation invariance and full functional relationship if and only if the value 1 is attained. The Linfoot correlation equals the absolute value of the Pearson correlation in case of bivariate gaussian random variables. For non-Gaussian data highly dependent random variables might have Pearson correlation zero, while the Linfoot correlation will be positive. However, being a transformation of the mutual information (equivalently the Kullback-Leibler divergence), the Linfoot measure is in general not straightforwardly estimable. Several estimators of mutual information have been suggested and their performances were evaluated in a variety of scenarios. Generally, bias is hard to control in this task and several alternative measures to gauge dependency have been suggested (eg., Murrel's MIC), but the alternatives do not fulfil the Renyi properties. We propose to estimate Linfoot correlation via a feedforward neural network on scenarios of increasing generality. By expanding from bivariate Gaussianity into more elaborate distributions we expect that estimators are learned having the potential of reflecting Linfoot correlation properties. We derive the Linfoot correlation of the Gaussian – and the Clayton copulas. The first is parametrized by the Pearson correlation while for the second a closed formula for the Linfoot correlation is derived, to our knowledge not seen before. We compare by simulation our estimator with other proposed ones.

Incorporating Memory into Continuous-Time Spatial Capture-Recapture Models

***Clara Panchaud**, University of Edinburgh*

Estimating wildlife species abundance underpins the conservation and management of animal populations and natural reserves. Capture-recapture studies involving repeated surveys are used to collect data on uniquely identifiable individuals through remote sensors, from which population estimates can be obtained through spatially explicit capture-recapture (SCR) models. SCR models consider spatial correlation by assuming that animals are more likely to be observed by sensors close to their activity centre. However, these traditional SCR models rely on the assumption that the probability an individual is observed at a given trap depends solely on the (unobserved) spatial location of their activity centre. This assumption implies that an individual's previous known location does not influence the probability of being seen at future times at the given trap locations. This implication is ecologically unrealistic given that animals move through space and time smoothly. We present a new continuous-time modeling framework to account for spatial correlation of observations due to both an individual's (latent) activity centre and (known) observed locations from previous captures. We consider observations as generated by an inhomogeneous temporal Poisson process and make use of the Ornstein-Uhlenbeck process, commonly used to model animal movement. We show that ignoring movement can lead to biased estimates and we present results from simulations and from a dataset of American martens.

Gaussian process regression for value-censored functional and longitudinal data

***Andreas Kryger Jensen**, University of Copenhagen, Denmark; Adam Gorm Hoffmann, University of Copenhagen, Denmark; Claus Thorn Ekstrøm, University of Copenhagen, Denmark; Benjamin Zeymer Christoffersen, Division of Robotics, Perception and Learning, KTH Royal Institute of Technology, Sweden and Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden; Andreas Kryger Jensen, University of Copenhagen, Denmark*

Gaussian process (GP) regression is widely used for flexible and non-parametric Bayesian modeling of data arising from underlying smooth functions. This paper introduces a solution to GP regression when the observations are subject to value-based censoring. We derive exact and closed-form expressions for the conditional posterior distributions of the underlying functions in both the single-curve fitting case and in the case of a hierarchical model where multiple functions are modeled simultaneously. Our method can accommodate left, right, and interval censoring, and is directly applicable as an empirical Bayes method or integrated in a Markov-Chain Monte Carlo sampler for full posterior inference. Our method is validated through extensive simulations, where it substantially outperforms naive approaches that either exclude censored observations or treat them as fully observed values. We give an application to a real-world dataset of longitudinal HIV-1 RNA measurements, where the observations are subject to left censoring due to a detection limit.

False discovery rate control via Bayesian mirror statistic

Marco Molinari, *Oslo Centre for Biostatistics and Epidemiology (OCBE), University of Oslo; Magne Thoresen, Oslo Centre for Biostatistics (OCBE), University of Oslo*

Simultaneously performing variable selection and inference in high-dimensional models is an open challenge in statistics and machine learning. The increasing availability of vast amounts of variables requires the adoption of specific statistical procedures to accurately select the most important predictors in a high-dimensional space, while being able to control some form of selection error. In this work we adapt the Mirror Statistic approach to False Discovery Rate (FDR) control into a Bayesian modelling framework. The Mirror Statistic, developed in the classic frequentist statistical framework, is a flexible method to control FDR, which only requires mild model assumptions, but requires two sets of independent regression coefficient estimates, usually obtained after splitting the original dataset. Here we propose to rely on a Bayesian formulation of the model and use the posterior distributions of the coefficients of interest to build the Mirror Statistic and effectively control the global FDR without the need to split the data. Moreover, the method is very flexible since it can be used with continuous and discrete outcomes and more complex predictors, such as with mixed models. We keep the approach scalable to high-dimensions by relying on Automatic Differentiation Variational Inference and fully continuous prior choices.

Estimating lead time in cancer screening programmes based on incidence comparison

Maja Pohar Perme, *University of Ljubljana, Medical Faculty, Department of Biostatistics and Medical Informatics; Bor Vratanar*

In cancer screening programmes, participants are regularly screened every few years in the attempt to detect early signs of cancer. Without screening, cancer would likely progress undetected until symptoms appear. The interval between early detection and the eventual onset of symptoms, had screening not been conducted, is known as lead time. Estimating lead time is challenging for two reasons. First, it is hypothetical— there is no direct way to measure when symptoms "would have" occurred for screen-detected cases as treatment starts right after screen diagnosis. Second, some screen-detected cancers are overdiagnosed: they would have died due to other causes before hypothetical symptoms take place. We propose a new method that uses cancer incidence data to estimate lead time. When a new screening programme is introduced, incidence rates typically surge initially and shift to younger ages, then gradually decline compared to pre-screening levels. This shift in cancer incidence, stratified by age groups and calendar years, carry vital information about lead time. The proposed method estimates the parameters of a pre-specified distribution that best explains the observed shift in cancer incidence while accounting for overdiagnosis, using maximum likelihood estimation principles. The method requires data on both cancer cases invited to the programme as well as on cancer cases not invited to the programme. Our approach is flexible, allowing for the inclusion of additional covariates and accounting for non-progressive tumours. We validated our method through simulations and applied it to data from the Slovenian breast cancer screening programme.

Estimation of the largest mean Gaussian mixture component with population genetic applications

***Andreas Futschik**, Johannes Kepler University Linz*

We propose a new method to estimate the mixture component with the largest mean parameter when the data come from a Gaussian mixture model. We discuss some properties of the method and show that it has advantages compared to classical approaches of inference such as the EM algorithm when there is a large number of components. The method relies on inference for the truncated normal distribution. Our motivating application comes from population genetics, where the effective population size N_e is an important parameter when specifying null models. We first explain how N_e has usually been estimated. Then we show how our proposed method may be used to identify the neutral N_e.

Contributed session 3

A boosting model for monotonic degradation: First Hitting Time analysis via a homogeneous Gamma process

Clara Bertinelli Salucci, *University of Oslo; Azzeddine Bakdi, Corvus Energy; Ingrid Kristine Glad, University of Oslo; Bo Henry Lindqvist, Norwegian University of Science and Technology; Erik Vanem, DNV; Riccardo De Bin, University of Oslo*

First Hitting Time models offer an interesting framework for time-to-event analysis by representing the event as the first passage of an underlying stochastic process across a threshold. This is particularly attractive in settings where prior knowledge about the degradation trajectory is available or requires a better understanding. In contexts where degradation is irreversible — such as material fatigue or the progression of specific diseases — a monotonic process provides a natural and interpretable way to encode such knowledge. We propose a novel boosting algorithm for First Hitting Time models based on a homogeneous gamma process, which enforces monotonicity in the degradation path. The algorithm incrementally constructs the predictor by combining weak learners while preserving the stochastic properties of the underlying gamma process and the first hitting time formulation. Its iterative nature makes it particularly suitable for high-dimensional problems, where variable selection and regularisation are essential. We demonstrate the method's versatility and predictive accuracy on different datasets from engineering and biomedical domains, as well as in simulation studies. Results indicate that the proposed approach yields competitive performance compared to standard survival models. This work broadens the scope of First Hitting Time models, making them more accessible for complex and high-dimensional time-to-event data, and provides a practical tool for researchers and practitioners aiming to integrate process-driven modelling into survival analysis.

GPTCM: Generalized promotion time cure model to identify subclonal driver genes and improve cancer prognosis

Zhi Zhao, *University of Oslo; Fatih Kizilaslan, University of Oslo; Shixiong Wang, Akershus University Hospital; Manuela Zucknick, University of Oslo*

Single-cell technologies provide an unprecedented opportunity for dissecting the interplay between the cancer cells and the associated tumor microenvironment, and the produced high-dimensional omics data should also augment existing survival modeling approaches. However, there is no statistical model to integrate multiscale data including individual-level survival data, multicellular-level cell composition data and cellular-level single-cell omics covariates. We propose a generalized promotion time cure model (GPTCM) for the multiscale data integration to identify subclonal driver genes for cancer prognosis. We demonstrate with simulations that our model is able to identify cell-type-associated covariates and improve survival prediction. We apply the model in a case study with nodal B-cell non-Hodgkin lymphoma patient data, where cancer cells are differentiated from various subtypes of B cells.

A comprehensive simulation study evaluating the predictive performance of Cox proportional hazards model and machine learning methods for time-to-event data

Josline Adhiambo Otieno, *Department of Statistics, Umeå University; Jenny Häggström, Department of Statistics, Umeå University; Marie Eriksson, Department of Statistics, Umeå University*

Many data-driven risk prediction models have been developed for analysing time-to-event data. However, choosing the most suitable model for accurate predictions in a specific medical application remains a challenge. Simulation enables effective comparison based on equal-sized datasets. This study provided a comprehensive and fair comparison of the survival prediction performance of machine learning (ML) models and the traditional Cox proportional hazards (PH) model, using both simulated and real datasets. ML models included random survival forests, extreme Gradient Boosting, and deep neural networks (DeepSurv). We

evaluated the performance of the models using the C-index and the Integrated Brier Score. The evaluation was performed under different data-generating mechanisms, such as varying sample sizes, censoring proportions, the addition of noise variables, and in the presence of different types of model misspecification. All the models showed improved predictive performance with increasing sample sizes. However, their performance declined as the proportion of censoring increased. An increase in noise variables reduced prediction accuracy across all models, regardless of dataset size. Tree-based models demonstrated promising predictive performance compared to the Cox PH model and DeepSurv in the presence of misspecification and a large number of noise variables. The Cox PH model performed well with larger sample sizes and fewer noise variables. It also performed well when the model was correctly specified or had only minor misspecification.

Investigating the effect of rare CYP2C19 and CYP2D6 genetic variants using population pharmacokinetic models in the Estonian Biobank

***Laura Birgit Luitva**, Institute of Mathematics and Statistics, University of Tartu, Tartu, Estonia; Hiie Soeorg, Institute of Biomedicine and Translational Medicine, University of Tartu; Märt Möls, Institute of Mathematics and Statistics, University of Tartu; Maris Alver, Estonian Genome Centre, Institute of Genomics, University of Tartu; Lili Milani, Estonian Genome Centre, Institute of Genomics, University of Tartu; Krista Fischer, Institute of Mathematics and Statistics, University of Tartu, Estonian Genome Center, Institute of Genomics, University of Tartu*

CYP2C19 and CYP2D6 are essential enzymes in the metabolism of various drugs. Genetic variants in CYP2C19 and CYP2D6 affect their activity and thereby drug metabolism. The effects of common genetic variants are well-known, however characterizing rare genetic variants remains challenging. We aim to investigate the effect of CYP2C19 and CYP2D6 genetic variants and other factors on drug metabolism by analyzing pharmacokinetic profiles of 114 individuals using population pharmacokinetic models. The pharmacokinetic profiles were obtained from a clinical study conducted in the Estonian Biobank where eligible study participants were orally administered omeprazole and metoprolol – drugs primarily metabolized by CYP2C19 and CYP2D6, respectively. The drug and metabolite concentration measurements were available from 10 timepoints from baseline to 8 hours after drug administration. The preliminary analysis using a basic approach of comparing the drug/metabolite area-under-the-curve ratios generally confirmed the effects of the characterized genetic variants and bioinformatically predicted effects of rare genetic variants. Currently, we are implementing population pharmacokinetic models to more accurately analyze the pharmacokinetic data and study the effect of CYP2C19 and CYP2D6 genetic variants and other factors on drug metabolism. Population pharmacokinetic models enable to estimate important pharmacokinetic parameters such as clearance and volume of distribution, and investigate the sources of variability. Compartmental modelling and differential equations are used for describing the absorption, distribution, metabolism, and elimination of the drug in the body. Statistical modelling involves non-linear mixed-effects models to describe the variability and impact of covariates on pharmacokinetic parameters.

Exploring evolutionary pathways of antimicrobial resistance genes through hypercubic clustering

***Kazeem Adesina Dauda**, University of Bergen; Iain Johnston, University of Bergen*

In recent years, genomic technology has been widely used to provide a far more detailed picture of the evolution and spread of antimicrobial resistance (AMR), resulting in large-scale genomic data- a major challenge. Using large-scale genomic data to learn and predict AMR evolutionary pathways is a scientific necessity and a matter of significant public interest, as it directly impacts global public health. The large-scale genomic data can be perceived as a sequential stochastic acquisition of binary traits, involving the presence or absence of genes. Therefore, to effectively address the increasing threat of AMR, it is imperative to understand how bacteria acquire AMR genes from large-scale genomic data through evolutionary dynamics. Previous studies explored the Hypercubic inference method based on a hidden Markov chain model to dynamically infer the acquisition of binary traits through the Bayesian method (HyperTraPS) and adapted Baum–Welch (expectation–maximization) algorithm (HyperHMM). However, these studies lack the capability of directly handling

large-scale genomic data. To overcome this barrier, we develop a flexible approach that combines the power of the clustering approach and the HyperHMM inference method called Cluster-HyperHMM. The study applies the proposed methods to the synthetic and real-life data on Klebsiella Pneumonia to compare the resistance patterns and evolutionary progressive pathways in the six continents. The results of evolutionary pathways revealed differences in the genome feature acquisition in various continents.

Joint probability approach for prognostic prediction of conditional outcomes: application to quality of life in head and neck cancer survivors

***Marissa LeBlanc**, University of Oslo, Norwegian Insitute of Public Health (FHI); Mauricio Moreira-Soares and Erlend I. F. Fossen, Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Aritz Bilbao-Jayo and Aitor Almeida, DeustoTech, University of Deusto, Bilbao, Spain; Laura Lopez-Perez and Itziar Alonso and Maria Fernanda Cabrera-Umpierrez and Giuseppe Fico, Life Supporting Technologies Research Group, ETSIT, Technical University of Madrid, Spain; Susanne Singer, Institute of Medical Biostatistics, Epidemiology and Informatics and University Cancer Center at University Medical Center of Johannes Gutenberg, University Mainz, Germany; Katherine J. Taylor, Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center of Johannes Gutenberg, University Mainz, Germany; Andrew Ness and Steve Thomas and Miranda Pring, Bristol Dental School, United Kingdom; Lisa Licitra, Head and Neck Medical Oncology Department, IRCCS National Cancer Institute Foundation, Milan, Italy; Stefano Cavalieri, Department of Oncology and Hemato-oncology, University of Milan, Italy; Arnoldo Frigessi, Oslo Centre for Biostatistics and Epidemiology, University of Oslo*

Background: Conditional outcomes are outcomes defined only under specific circumstances. For example, future quality of life can only be ascertained when subjects are alive. In prognostic models involving conditional outcomes, a choice must be made on the precise target of prediction: one could target future quality of life, given that the individual is still alive (conditional) or target future quality of life jointly with the event of being alive (unconditional). We aim to (1) introduce a probabilistic framework for prognostic models for conditional outcomes, and (2) apply this framework to develop a prognostic model for quality of life 3 years after diagnosis in head and neck cancer patients.
Methods: A joint probability framework was proposed for prognostic model development for a conditional outcome dependent on a post-baseline variable. Joint probability was estimated with conformal estimators. We included head and neck cancer patients alive with no evidence of disease 12 months after diagnosis from the UK-based Head & Neck 5000 cohort ($N = 3572$) and made predictions 3 years after diagnosis. Predictors included clinical and demographic characteristics and longitudinal measurements of quality of life. External validation was performed in studies from Italy and Germany.
Findings: Of 3572 subjects, 400 (11.2%) were deceased by the time of prediction. Model performance was assessed for prediction of quality of life, both conditionally and jointly with survival. C-statistics ranged from 0.66 to 0.80 in internal and external validation, and the calibration curves showed reasonable calibration in external validation. An API and dashboard were developed.
Interpretation: Our probabilistic framework for conditional outcomes provides both joint and conditional predictions and thus the flexibility needed to answer different clinical questions. Our model had reasonable performance in external validation and has potential as a tool in long-term follow-up of quality of life in head and neck cancer patients.

Contributed session 4

Occupational exposures and kidney cancer among 25 000 male offshore petroleum industry workers: relative risks and healthy worker survivor bias

***Nita Kaupang Shala**, Department of Research, Cancer Registry of Norway, Norwegian Institute of Public Health & Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Marit B Veierød, Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Ronnie Babigumira, Cancer Registry of Norway, Norwegian Institute of Public Health & Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Leon AM Berge, Cancer Registry of Norway, Norwegian Institute of Public Health & Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Sven O Samuelsen, Department of Mathematics, University of Oslo; Jorunn, Kirkeleit, Department of Occupational Medicine and Epidemiology, National Institute of Occupational Health, Oslo, Norway; Magne Bråtveit, Department of Global Public Health and Primary Care, University of Bergen; Melissa C Friesen and Alexander P Keil and Debra T Silverman and Nathaniel Rothman and Lan Qing, Division of Cancer Epidemiology and Genetics, Occupational and Environmental Epidemiology Branch, National Cancer Institute, Bethesda, MD, USA; Jo S Stenehjem, Cancer Registry of Norway, Norwegian Institute of Public Health & Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Tom K. Grimsrud, Cancer Registry of Norway, Norwegian Institute of Public Health*

Background: Kidney cancer has been a suspected occupational disease in petroleum workers. Health conditions that are linked to kidney cancer may prompt termination or change of work, and thereby restrict occupational exposures in high-risk individuals, creating a healthy worker survivor bias (HWSB). This bias, a form of time-varying confounding, presents significant challenge to fully understand the health impacts of various exposures.
Methods: We examined associations between occupational exposures and kidney cancer among males in the Norwegian Offshore Petroleum Workers (NOPW) cohort using a case-cohort design, with 169 incident cancers identified by linkage to national registry data (1999–2021) and a subcohort of 2090 non-cases, all employed 1965–1998. Relative risks (hazard ratios, HRs) by cumulative exposure to benzene, crude oil, chlorinated degreasing agents (CDA), or surface treatment (priming, painting), were estimated by weighted Cox regression.
Results: Inverse exposure-response trends suggested HWSB, reinforced by analyses of necessary components of HWSB. Bias was partly alleviated by adjustment for total employment duration, and by 20-year lagging of cumulative exposure to benzene, crude oil, or CDA. Workers in surface treatment (ever vs. never) showed increased HR=2.22, 95% confidence interval 1.04–4.72 (9 cases only).
Conclusions: Based on our analysis we could neither confirm nor exclude an occupational impact on kidney cancer. Depending on the agent, there were indications that more advanced statistical methods could be useful in avoiding residual time-varying confounding.

Prediction of long-term cumulative cancer incidence by lifestyle risk factors in Finland

***Markus Linjamäki**, Finnish Cancer Registry; Karri Seppä, Finnish Cancer Registry*

Background: Reliable forecasts of cancer incidence are crucial for effective public health planning. The cause-specific cumulative incidence function takes account the competing risk of death, giving the estimate of probability of developing cancer. Our aim is to predict cumulative cancer incidence and to quantify the impact of the major lifestyle risk factors on cumulative incidence.
Methods: A Bayesian age-period-cohort model was fitted to the pooled data of seven health studies conducted in Finland between 1972 and 2015 with 224,820 participants and 20,253 first primary cancers diagnosed during follow-up. A generalized additive model with a tensor product was used to model the effects of age, period and cohort and the lifestyle factors (smoking status, alcohol consumption and body mass index) on the cause specific hazard of cancer incidence and competing mortality, respectively. We estimated the cumulative incidence function of first cancer for each combination of the lifestyle factors by birth cohort.
Results: Preliminary results show that among current smokers, the probability of developing cancer between ages 40 and 90 was 54% (95% credible interval: 46–62%) for men born in 1960–1964, and 49% (95% CI:

39–61%) for women. Among never smokers with other lifestyle factors similar to current smokers, the probability was 46% (95% CI: 38–56%) for men and 41% (95% CI: 31–53%) for women.
Conclusions: The presented model allows the estimation of long-term cumulative incidence with appropriate estimates of uncertainty.

Associations between lifestyle factors and cancer survival in Finland

***Karri Seppä**, Finnish Cancer Registry; Iitu Peltola and Sanna Heikkinen, Finnish Cancer Registry; Johan G Eriksson and Tommi Härkänen and Pekka Jousilahti and Paul Knekt and Seppo Koskinen and Satu Männistö, Finnish Institute for Health and Welfare; Ossi Rahkonen, University of Helsinki; Nea Malila, Finnish Cancer Registry; Maarit Laaksonen, Finnish Institute for Health and Welfare; Janne Pitkäniemi, Finnish Cancer Registry*

Background: Modifiable risk factors, such as smoking, alcohol consumption, and excess weight, are associated with cancer incidence but evidence of their role in cancer survival is limited.
Material and methods: Pooled data from seven health studies conducted in Finland during 1972-2015 were used to assess the associations of smoking, alcohol consumption, body mass index, physical inactivity, and education on mortality among 224,820 participants, of whom 10,637 were diagnosed with colorectal, lung, pancreatic, prostate or breast cancer. We compared patients' mortality to that of a cancer-free population and estimated relative excess risks of death (RER) using piecewise constant excess hazard models.
Results: Current smoking was associated with an increased excess mortality in male pancreatic cancer (RER compared to never-smokers 1.54, 95% CI 1.08-2.19) and prostate cancer patients (1.55, CI 1.08-2.23), as well as in female lung cancer patients (1.44, CI 1.07-1.92). Obesity was associated with an increased excess mortality in prostate cancer patients (1.56, CI 1.06-2.30) and physical inactivity in colorectal cancer (1.33, CI 1.01-1.76 in men; 1.36, CI 1.06-1.74 in women) and male lung cancer patients (1.15, CI 1.02-1.30). Excess mortality was consistently elevated among patients with low education.
Conclusions: Several risk factors of cancer were associated with increased excess mortality among cancer patients. Mortality among patients with low education remained elevated even after adjusting for the cancer stage and the lifestyle factors. It is notable that even in a Nordic welfare state with potentially equal health care, marked socioeconomic inequalities persist in mortality among cancer patients.

Contributed session 5

Parametric estimation and comparison of age-reading error matrices for fish stock assessments

***Solveig Engebretsen**, Norwegian Computing Center; Magne Aldrin, Norwegian Computing Center; Florian Berg, Norwegian Institute of Marine Research*

Stock assessments of fish are based on age-structured data from commercial catches and scientific surveys. The age estimates are based on expert readers interpreting calcified structures, typically scales or otoliths. As the readability of the calcified structures for individual fish may vary, and due to human error and subjectivity, the read age may deviate from the true age of the fish. In samples from commercial catches and scientific surveys, each individual fish is typically only read once. However, reading exchange events, where multiple readers have estimated the age of each individual fish, are regularly hosted with the purpose of training readers and to quantify the age-reading errors. In this work, we propose a parametric model for age-reading error matrices. The parameters allow for comparing properties of these matrices for e.g. different stocks, species, and calcified structures. One major drawback with such age-reading exchanges is that the true age of the fish is unknown. The age comparison guidelines recommended by, e.g. the International Council for the Exploration of the Sea, suggest to use the modal age among the readers as a proxy for the true age. We assess the effect of this assumption through simulation. While most stock assessment models can take into account age-reading errors, these have to be provided as an input to the models. We also assess the effect of accounting for age-reading errors in stock assessment.

Cleaner fish performance in Norwegian Salmonid Farms

***Nora Røhnebæk Aasen**, Norwegian Computing Center; Solveig Engebretsen, Norwegian Computing Center; Magne Aldrin, Norwegian Computing Center; Peder Jansen, Aqualife R&D*

Salmon lice is a main concern for fish farmers in Norway, with consequences for the welfare of both farmed and wild salmon. It is also the main limiting factor for increasing the production of farmed salmon in Norway. As one of several control measures against salmon lice, cleaner fish, which prey on salmon lice attached to the salmon, is a commonly used preventative measure. However, there are few quantitative estimates of the effect of cleaner fish. Recently, several farmers have stopped deploying cleaner fish due to uncertainty around their efficiency and welfare concerns. In this work, we have used a partly stage-structured model for lice abundance on salmonid farms along the Norwegian coast to estimate the effect from cleaner fish as a preventative lice measure. The dataset used in the analysis is a combination of openly available data from BarentsWatch on all Norwegian salmon farms, and data from the Norwegian Directorate of Fisheries. The cleaner fish specification in the model is based on the Hollings Functional Response, which means that consumed lice is modelled as a function of the lice density. We also let one parameter depend on temperature, and we find different lice grazing efficiency under different temperatures. The model distinguishes between two cleaner fish species: lumpfish and wrasse (but not subspecies of wrasse). We also estimate the effect for adult female lice and other motile lice (adult male lice and pre-adult stages of salmon lice) separately.

Combining different treatments to delay resistance in controlling salmon lice in fish farms

***Ragnar Bang Huseby**, Norwegian Computing Center; Magne Aldrin, Norwegian Computing Center*

We investigate various approaches to reduce the problem of salmon louse for the salmon farming industry. Salmon lice are parasites that can be spread through water from one fish farm to wild fish and to neighbouring fish farms. Various measures have been used to control salmon lice in fish farms, but frequent use of the same type of treatment may cause resistance, especially for medicinal treatments and possibly for non-medicinal treatments (e.g. freshwater). Our work focuses on resistance development caused by genetic selection. This means that some genotypes of lice have a higher probability of surviving a given treatment than other lice, and consequently the proportion of lice of resistant genotypes tends to increase by using that kind of

treatment. We consider treatment policies involving two types of treatments with similar efficacy properties, but with independent selection mechanisms. The investigation is done by simulating the population of lice in neighbouring fish farms using a population model for lice. When the density of lice on a small sample of fish exceeds a given threshold, a treatment, which type depends on the strategy, is applied. For each simulation we record the number of required treatments and the levels of resistance. Regarding delaying resistance our results suggest that applying the two treatment types in combination is more effective than any of the considered strategies with separate use of the two types. This is joint work with researchers at the Norwegian Veterinary Institute, Aqualife R&D, Pharmac and the Norwegian University of Life Sciences.

Contributed session 6

Quantification of vaccine waning as a challenge effect

***Matias Janvin**, University of Oslo; Mats J. Stensrud, Federal Polytechnic School of Lausanne (EPFL)*

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

Long-term effect of pharmacological treatment on academic achievement of Norwegian children diagnosed with ADHD: a target trial emulation

***Tomás Varnet Pérez**, Norwegian Institute of Public Health & University of Oslo; Kristin Romvig Øvergaard, University of Oslo & Oslo University Hospital; Arnoldo Frigessi, Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Guido Biele, Norwegian Institute of Public Health & University of Oslo*

Background: Attention-deficit/hyperactivity disorder (ADHD) is one of the most commonly diagnosed mental disorders in children. For many patients, treatment involves long-term medication in order to reduce symptoms, regulate behaviour, and, hopefully, improve school performance and achievement. However, there is little to no evidence to support a long-term effect on the latter complex outcomes.
Methods: We utilize a target trial framework to emulate a pretest-posttest control group design and estimate the intention-to-treat effect of ADHD medication on national test scores in children diagnosed with ADHD born between 2000 and 2007 in Norway. The data was obtained through linkage of Norwegian registries. Sensitivity analyses include a negative control exposure, proximal causal inference, and a post-hoc informal replication study of a conflicting previous study.
Results: The resulting analytic sample size consisted of 8 548 children diagnosed with ADHD, with about 9% missingness in their grade eight national test scores. We find that initiating ADHD medication had a slight positive average effect on national test scores for all three domains: English, numeracy and reading (standardized mean differences: 0.037 (95%-compatibility interval (CI) = [-0.003 ; 0.076]), 0.063 (95% CI = [0.016 ; 0.111]), 0.071 (95% CI = [0.030 ; 0.111]) respectively).
Conclusion: We conclude that the estimated long-term average effect of ADHD medication on learning, as measured by the Norwegian national tests, is not clinically relevant. Study strengths include the use of real-world data on ecologically valid and relevant outcomes, the robustness of results across model specifications. Limitations include possibility of unobserved confounding and lack of prescription data.

Estimating opioid-sparing treatment effects using natural treatment values

***Catharina Stoltenberg**, University of Oslo; Matias Janvin, University of Oslo; Leiv Arne Rosseland, University of Oslo; Jon Michael Gran, University of Oslo*

There is a need to identify alternative treatment regimes that reduce opioid reliance without compromising pain management. One promising approach is to combine opioid-sparing medications with opioids. These medications reduce required opioid doses while maintaining effective pain relief, thereby reducing the harmful side effects associated with opioids. A wide range of medications is being investigated for their analgesic and opioid-sparing potential. However, the evidence supporting their actual opioid-sparing effects varies

considerably. We present a formal framework for estimating effects of opioid-sparing regimes on subsequent opioid use in observational and experimental data. Specifically, we perform causal inference by emulating or conducting a randomized experiment comparing two treatment strategies: administering a presumed opioid-sparing medication, such as nonsteroidal anti-inflammatory drugs (NSAIDs), whenever opioids are administered during a given treatment period versus not administering NSAIDs at those same time points. When opioids are not administered, NSAIDs may be used as intended outside of the regime. This class of treatment regimes, referred to as add-on regimes, leverages natural opioid and NSAID values to provide treatment strategies that align with clinical practice. We derive identification and estimation results for the expected opioid dose under add-on regimes, consistent with established results. These proofs rely on assumptions of no unmeasured confounding, which involves fewer variables than those used in established results, making it easier for practitioners to assess their validity. We apply the theoretical results to estimate the opioid-sparing effect of NSAIDs using data from a cohort of Norwegian trauma patients, linking the Norwegian National Trauma Registry to several national databases. Overall, our analyses suggest that NSAIDs have an opioid-sparing effect.

Definition, identification, and estimation of the direct and indirect number needed to treat

***Valentin Vancak**, Holon Institute of Technology; Arvid Sjölander, Karolinska Institutet*

The number needed to treat (NNT) is an efficacy and effect size measure commonly used in epidemiological studies and meta-analyses. The NNT was originally defined as the average number of patients needed to be treated to observe one less adverse effect. In this study, we introduce the novel direct and indirect number needed to treat (DNNT and INNT, respectively). The DNNT and the INNT are efficacy measures defined as the average number of patients that needed to be treated to benefit from the treatment's direct and indirect effects, respectively. We start by formally defining these measures using nested potential outcomes. Next, we formulate the conditions for the identification of the DNNT and INNT, as well as for the direct and indirect number needed to expose (DNNE and INNE, respectively) and the direct and indirect exposure impact number (DEIN and IEIN, respectively) in observational studies. Next, we present an estimation method with two analytical examples. A corresponding simulation study follows the examples. The simulation study illustrates that the estimators of the novel indices are consistent, and their confidence intervals meet the nominal coverage rates.

Using directed acyclic graphs to determine whether multiple imputation or subsample multiple imputation estimates of an exposure-outcome association are unbiased

***Paul Madley-Dowd**, University of Bristol; Rachael A. Hughes, University of Bristol; Maya B. Mathur, Stanford University; Jon Heron, University of Bristol; Kate Tilling, University of Bristol*

Missing data is a pervasive problem in epidemiology, with multiple imputation (MI) a commonly used analysis method. MI is valid when data are missing at random (MAR). However, definitions of MAR with multiple incomplete variables are not easily interpretable and graphical model-based conditions are not accessible to applied researchers. Previous literature shows that MI may be valid in subsamples, even if not in the full dataset. Practical guidance on applying MI with multiple incomplete variables is lacking. We present an algorithm using directed acyclic graphs to determine when MI will estimate an exposure-outcome coefficient without bias. We extend the algorithm to assess whether MI in a subsample of the data, in which some variables are complete, and the remaining are imputed, will be valid and unbiased for the same coefficient. We apply the algorithm to several simple exemplars, and in a more complex real-life example highlight that only subsample MI of the outcome would be valid. Our algorithm provides researchers with the tools to decide whether (and how) to use MI in practice when there are multiple incomplete variables. Further work could focus on the likely size and direction of biases, and the impact of different missing data patterns.

Contributed session 7

A Julia implementation for estimating local ancestry in genetically admixed populations

Bjarke G Poulsen, *Center for Quantitative Genetics and Genomics, Aarhus University, Denmark; Huiming Liu and Doug Speed and Viktor Milkevych and Luc Janss and Emre Karaman, Center for Quantitative Genetics and Genomics, Aarhus University, Denmark*

Our project aims to develop an improved tool for determining ancestral origins of marker alleles in admixed individuals, i.e., local ancestry. State-of- the-art genomic prediction methods in genetically admixed populations rely on accurate estimation of local ancestry. Generally, software for estimating local ancestry divides the genome into fixed-size windows based on either physical distances or numbers of markers. However, this approach may be not be advantageous because genomic regions provide different amounts of information for ancestry assignment. Therefore, we propose a novel method for estimating local ancestry using a Naive Bayes classifier that allows for flexible window sizes. We have investigated both the accuracy and run-time using simulated data for five chromosomes (total length 670 centiMorgans with on average 2120 markers per chromosome). The simulated data consisted of three ancestral populations and one admixed population. The first generation of the admixed population was created by mating females from ancestral population A with males from ancestral population B. Then, generations two to four were created by mating females from the previous generation in the admixed population with males from population C, A, and B, respectively. The simulation was repeated 10 times. Preliminary results showed that the accuracy was 99% for the local ancestries of maternally inherited alleles among admixed individuals from generation four, and that the run-time of the package was competitive with existing software tools such as RFMIX.

Accounting for uncertainty in residual variances improves calibration of the Sum of Single Effects model for small sample sizes

William R.P. Denault, *Department of Statistics and Department of Human genetics, University of Chicago & Oslo Centre for Biostatistics and Epidemiology (OCBE), Oslo University Hospital*

The Sum of Single Effects (SuSiE) model, and a fitting procedure based on variational approximation, were recently introduced as a way to capture uncertainty in variable selection in multiple linear regression. This approach is particularly well-suited to cases where variables are highly correlated, and it has become quickly adopted for genetic fine-mapping studies. Here, we show that in small studies (<50 samples) the original fitting procedure can produce substantially higher rates of false positive findings than it does in larger studies. We show that a simple alternative fitting procedure, which takes account of uncertainty in the residual variance, notably improves performance in small sample studies.

PliableBVS: A flexible Bayesian variable selection method for modeling interactions with mandatory modifying variables

Theophilus Quachie Asenso, *Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Manuela Zucknick, Oslo Centre for Biostatistics and Epidemiology, University of Oslo*

Modelling interactions in high-dimensional data is a notoriously difficult problem. Analyzing high-dimensional data with conventional tools is very challenging. One of the reasons is that most existing models cannot easily handle cases with high-dimensional data and many interaction effects among the covariates, and they often make strong assumptions, e.g. strong hierarchy between main and interaction effects. Another reason is the challenge in the reduction of false positves of features when performing variable selction during, for example, biomaker discoveries. In this work we propose a flexible Bayesian variable selection (BVS) method to estimate interaction effects with motivation from the pliable lasso model which assumes an asymmetric weak hierarchy when including interactions. We also demonstrate how the method can be used to estimate interactions. Our proposed method is a BVS model with spike-and-slab prior where the asymmetric hierarchy between interactions and main effects is implemented via the inclusion indicator variable. We will

present the model "PliableBVS" and show how effective it is in handling interactions. We will also show the performance of the model in making predictions and reducing false positive rate, through data simulations and real data applications. In two real data applications, we will show results from using metabolomics and proteomics data measured at repeated time points in pregnant women, either to predict the time to onset of labor (example 1), or to predict late-onset preeclampsia (example 2).

Integrating multiple data sources with interactions in multi-omics using cooperative learning

***Matteo D'Alessandro**, Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Theophilus Quachie Asenso, Oslo Centre for Biostatistics and Epidemiology, University of Oslo; Manuela Zucknick, Oslo Centre for Biostatistics and Epidemiology, University of Oslo*

Multiomics data integration is a powerful approach to understanding complex biological processes and guiding personalized medicine. However, patient groups are heterogeneous, influencing how biological processes affect clinical outcomes. One might therefore wish to model some of this heterogeneity by including potential interactions with modifying variables, e.g. disease subtype, genetic modifiers, patient age and gender, or time. This study aims to achieve this goal by extending a hierarchical interaction model, the pliable lasso, to incorporate multi-source data. The pliable lasso models interactions between high-dimensional predictors (e.g., omics data) and a predefined set of modifying variables. Cooperative learning integrates multiple data sources by introducing an "agreement penalty" in the objective function, encouraging predictions from different omics platforms to align. The developed model, which combines these two frameworks, was evaluated on simulated datasets and two real- world applications: predicting labor onset and cancer treatment response. Our approach demonstrated superior predictive performance compared to existing methods, particularly when data sources shared underlying biological signals. In the labor onset dataset, the model identified key proteomic and metabolomic markers of gestational timing. In the cancer dataset, it revealed significant interactions between cancer tissue types and genomic features relevant to treatment response. The extension of pliable lasso to a cooperative learning framework addresses critical challenges in multiomics data analysis and allows flexible interaction modeling and multi-source integration. This approach is especially suited to biological settings where a set of modifying variables influences the relationship between omics variables and the response. The paper is available in preprint at https://arxiv.org/abs/2409.07125.

Estimation of risk ratios and bias-adjusted risk ratios from logistic regression models by transforming predicted probabilities into beta distributions

***Esben Østergaard Eriksen**, Department of Production Animal Clinical Sciences, Norwegian University of Life Sciences; Merete Forseth, Norsk Kylling AS*

Logistic regression models are often used in the analysis of epidemiological studies of both animal and human health. Exponentiation of coefficients from these models yields odds ratios; however, risk ratios (RR) are desirable due to the ease of interpretation. When bias is not sufficiently adjusted for in causal inference studies, it is relevant to conduct a quantitative bias analysis (QBA). Some straightforward implementations of QBA require a 2x2 contingency table as an input to estimate bias-adjusted RRs representing causal effects. This presentation shows how to calculate the parameters of beta distributions (α, β) from the mean estimate and confidence interval of marginal or conditional probabilities predicted from logistic models. Since beta distribution parameters represent the outcomes of a series of Bernoulli trials (n positives + 1 and n negatives + 1), it is proposed that the parameter values (minus 1) can be used as the cell frequencies in a 2x2 contingency table. This allows for the calculation of the RR with confidence intervals with established methods and the conduction of QBA. Satisfactory coverage of such confidence intervals is demonstrated in a simulation study. An applied example is shown from a study fitting a hierarchical multivariable multinomial logistic model to estimate the causal effect of broiler chicken hybrid on the risk of foot pad dermatitis with adjustment for an insufficient set of confounders. Conditional predicted probabilities are obtained from the model, and the proposed method is used to estimate the RR and the RR adjusted for an unmeasured confounder using QBA.

Poster session

Are genetic changes in 15q13.3 associated with lower IQ score?

***Audrone Jakaitiene**, Vilnius University; Tadas Žvirblis, Vilnius University; Pilar Caro, Heidelberg University; Christian P. Schaaf, Heidelberg University*

Genetic alterations in the 15q13.3 region are associated with rare neurodevelopmental disorders, such as autism, epilepsy, and schizophrenia. This area contains around ten genes, and treatments mainly target symptoms rather than the root causes. Not all individuals with 15q13.3 copy number changes exhibit symptoms, making it challenging to predict severity and clinical outcomes, which complicates modeling health impacts. This multi-center prospective study evaluates cerebral activity and neural network function in individuals with 15q13.3 microdeletion or microduplication. We plan to enroll 15 subjects for each genetic variation and 15 healthy controls. Participants will undergo electrophysiological brain network analysis, IQ testing, and genetic assessments. The study protocol was approved by the Ethics Board of the Medical Faculty of Heidelberg University No. S-212-2023. An interim analysis identified ten subjects with genetic changes in chromosome 15q13.3; seven had a deletion, and three had a duplication. The mean age was 29.8 years (SD=13.36), with 30.0% being adolescents and 50.0% male. The mean IQ score was 80.1 (SD=17.30), significantly lower than the population average (p=0.006). Female subjects scored slightly higher than males, with means of 81.2 (SD=16.27) and 79.0 (SD=20.14), respectively. The interim analysis indicates that individuals with 15q13.3 microdeletion or microduplication have lower IQ scores than the average. This work is part of the EJP RD project "Resolving complex outcomes in 15q13.3 copy number variants using emerging diagnostic and biomarker tools (Resolve 15q13)" No. DLR 01GM2307 and has received funding from EJP RD partner the Research Council of Lithuania (LMTLT) under grant agreement No. S-EJPRD-23-1.

Statistical methods to assess the flow of brain fluid

***Are Hugo Pripp**, Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital & KG Jebsen Centre for Brain Fluid Research, University of Oslo; Geir Ringstad, Department of Radiology, Oslo University Hospital & KG Jebsen Centre for Brain Fluid Research, University of Oslo, Norway & Institute of Clinical Medicine, Faculty of Medicine, University of Oslo; Per Kristian Eide, Department of Neurosurgery, Oslo University Hospital, & KG Jebsen Centre for Brain Fluid Research, University of Oslo & Institute of Clinical Medicine, Faculty of Medicine, University of Oslo*

Cerebrospinal fluid (CSF) is a clear, colorless liquid that surrounds the central nervous system (CNS). Over the last few years, there has been a paradigm shift in our understanding of CSF passage in and around the CNS. A pivotal discovery is the brain-wide perivascular transport system for CSF, referred to as the glymphatic system, and enabling delivery of nutrient to the brain as well as efflux of waste products form brain metabolism. Our research group has been central in translating the discoveries originally made in rodents to humans. Assessing glymphatic function involves the use of tracers combined with magnetic resonance imaging (MRI) and advance mathematical processing. From a biostatistical perspective, evaluating the brain fluid flow and glymphatic activity presents challenges. Studies often involve small and heterogeneous patient samples, complex data management, large individual variation, clustered and repeated measurements, and non-linear relationships. Our research group has addressed these challenges using statistical methods ranging from simple descriptive statistics and hypothesis tests to linear mixed models, fractional polynomial regression, segmented (piecewise) regression and pharmacological modelling. We also employ visualization techniques that account for both group differences and patient variability. Our findings show that brain fluid flow is influenced by sleep and affected by neurological conditions such as Alzheimer's disease, brain injury, and stroke. Additionally, we discuss innovations in statistical assessment methods and their implications for clinical medicine and novel treatments for brain diseases.

How do municipal investments in elementary education modify the effects of parental unemployment on children's mental health in Sweden?

***Natalia Andreeva**, Umeå University; Anna Baranowska-Rataj, Centre for Demographic and Ageing Research, Umeå University; Xavier de Luna, Department of Statistics/USBE, Umeå University*

Parental job losses have a detrimental impact on the mental health of children (Baranowska-Rataj et al. 2024; Högberg & Baranowska-Rataj 2024). Previous literature has explored how the magnitude of the negative effects varies depending on the resources of families prior to job loss. However, little is known whether the effects of parental job losses vary across broader social environments such as municipalities. Local authorities have resources that may potentially compensate for the harmful effects of socioeconomic disadvantage among children. Identifying the buffering role of local authorities' resources advances the knowledge on the relevance of public institutions for population health. Our data comprise demographic covariates of approximately 70,000 children linked with their parents, with observations ranging from 2005 to 2013. We identified involuntary job losses of parents based on information on workplace closures. Measures of children's mental health outcomes were derived from the Prescribed Drug Register, which includes information about medicines prescribed for anxiety disorders and mood disorders - the most salient internalizing disorders among children in Sweden today. The municipal expenditures per child in elementary school were collected from the database "Politics, Institutions and Services in Swedish Municipalities" (Dahlström & Tyrberg 2016). With the causal forest method (Wager & Athey, 2018), designed to uncover heterogeneity of treatment effects, we computed the aggregated effects across municipalities and performed tests for the effect heterogeneity along the contextual variables representing the municipal investments. Results demonstrated a positive role of more resourceful municipalities in compensating for the negative impact of a job loss.

Gearing Statisticians up for Software Success

***Audrey Yeo Te-ying**, Finc Research*

Trial statisticians write software and are well placed to contribute to analytical solutions that inform business and clinical decision making. Early conversations about incorporating software engineering competence in a trials statistician's (Sabanés Bové, 2023) already impressive toolbox have started to emerge with the use of the common language of R. I share my personal experience about the actions and attitudes to achieve a state-of-art solution called `phase1b` and highlight good software engineering principles and ways of working. The package `phase1b` (Yeo et al, 2024) is a flexible toolkit that evaluates efficacy and futility analysis within the Bayesian framework for early Oncology trials. The R package informs decision making on whether the drug of concern warrants further investment. Since the evaluation of the efficacy and futility impacts decisions in life sciences research, it is advantageous that the profession strives to understand and create the conditions for more statistical software success such as the `phase1b`. Package available at https://genentech.github.io/phase1b.

Matched pairs with rare event outcomes

***Christiana Kartsonaki**, University of Oxford*

We consider the comparison of the rate at which a rare event occurs under different conditions which are present during two different time periods. For example, we consider a set of hospitals where the availability of a particular treatment is and is not available during a time period. The patient populations are broadly similar but the individual patients studied are distinct in different time periods. We are interested in the effect of the different conditions on the rate of occurrence of a rare event of interest. We assume that the number of events in each group has a Poisson distribution. The random variables (Y, Z) associated with a specific hospital are assumed to have a common value of an underlying rate μ multiplied by a factor representing the availability of the treatment. The nuisance parameters μ are eliminated from the likelihood by conditioning on the total number of events to yield a log likelihood that can be maximised to estimate the parameters representing the effect of the condition on the outcome. I would like to thank David R. Cox for his help with this work.

Assessing adverse events in hospitals with the Global Trigger Tool method; status quo and further developments

***Ingunn Fride Tvete**, Norwegian Computing Center; Ellen Catharina Tveter Deilkås, Norwegian Directorate of Health; Linda Reiersølmoen Neef, Norwegian Computing Center; Wenche Patrono, Norwegian Directorate of Health; Hanne Narbuvold, Norwegian Directorate of Health; Marion Haugen, The Norwegian Computing Center*

Objectives Global Trigger Tool (GTT) is a retrospective review method where trained GTT teams examine randomly sampled hospital records for detecting adverse events (AEs). The method assumes consistency in how teams conduct their record reviews for the hospitals to monitor their AE rate over time. For a possible extension of the GTT method to also do comparison of the AE rates across hospitals, it is of interest to consider several teams' agreement when rating the same records. This has not previously been explored. Methods We examined eight teams' reproducibility for rating the presence of at least one AE or not in 120 records per team, by comparing review results from two different ratings of these records. We further analyzed the agreement among five teams rating the same 200 records. We computed and compared several well-known agreement measures. Conclusions All over, we found the teams' ability to reproduce their previous findings to be good, indicating the method to work as intended. The agreement among the five teams rating the same 200 records was, on the other hand, moderate. Extending the GTT framework for comparisons across hospitals requires the method to be further developed. We discuss possible adjustments according to different patient compositions. We also suggest a stronger common platform of rating policies for a more unified way of conducting the review. These are novel analyses and bring new insight into the GTT review process.

A multi-center study on the consistency of drug response assays in AML patient-derived cells

***Katarina Willoch**, Department of Cancer Genetics, Oslo University Hospital (OUS) & Oslo Centre for Biostatistics and Epidemiology (OCBE); Zhi Zhao, Oslo Centre for Biostatistics and Epidemiology (OCBE); Pilar Ayuda-Durán, Institute for Cancer Research and Center for Cancer Cell Reprogramming, Institute of Clinical Medicine; Flora Mikaeloff, Department of Laboratory Medicine, Karolinska Institute; Tom Erkers, Department of Oncology-Pathology, Karolinska institute; Jorrit M. Enserink, Department of Molecular Cell Biology, Institute for Cancer Research and Center for Cancer Cell Reprogramming, Institute of Clinical Medicine and Section for Biochemistry and Molecular Biology, Faculty of Mathematics and Natural Sciences; Jeffrey W. Tyner, Oregon Health & Science University; Tero Aittokallio, Department of Cancer Genetics, Oslo University Hospital & Oslo Centre for Biostatistics and Epidemiology (OCBE) & Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki*

Ex vivo drug response assays have the potential to optimize treatment selection for personalized medicine. However, a widespread implementation of drug testing in patient cells is constrained by inconsistencies in reproducibility across experiments conducted in different laboratories. Here, we investigate the experimental variables contributing to drug response differences across four different labs using linear mixed models and data from four AML cohorts. We discovered that the type of positive controls used on drug plates, the number of concentrations the drugs were tested, and the viability assay used for response readout led to systematic differences in response levels. Moreover, the use of optical density for counting cells led to a significantly higher response levels, compared to using countess for counting cells. Other experimental factors, such as using MCM medium compared to HS-5 medium, did not result in significantly different responses. In conclusion, the ex vivo drug response assays could yield more robust and reproducible results if the experimental procedures were standardized and consistently applied across all laboratories.

A multistate model for waning of vaccine effectiveness and infection rates from cross sectional serology data for whooping cough

***Jonas Christoffer Lindstrøm**, Norwegian Institute of Public Health; Audun Aase and Gro Tunheim and Tove Karin Herstad, Norwegian Institute of Public Health*

We develop a model for antibody levels against Pertussis (i.e. whopping cough) after vaccination, which is used to analyze data from a cross- sectional seroprevalence study. The model is a variant of the catalytic model with waning and allows us to study both seroreversion (i.e. waning) and seroconversion (i.e. infection) rates. It thus helps us disentangle the effects of vaccination, waning, and infection on the antibody levels.