
27th Young Statisticians Meeting

YSM 27

PROGRAMME — ABSTRACTS — PARTICIPANTS

SEPTEMBER 29 - OCTOBER 1, 2023
OSIJEK, CROATIA



SUPPORTED BY

SCHOOL OF APPLIED MATHEMATICS AND INFORMATICS, UNIVERSITY OF OSIJEK
FOUNDATION OF THE CROATIAN ACADEMY OF SCIENCES AND ARTS
OSIJEK-BARANJA COUNTY AND CITY OF OSIJEK
CROATIAN BURO OF STATISTICS
IPSOS WEST ADRIA

Local Organizing Committee

Mirta Benšić, University of Osijek, Croatia
Ksenija Dumičić, University of Zagreb, Croatia
Danijel Grahovac, University of Osijek, Croatia
Ivan Papić, University of Osijek, Croatia
Kristian Sabo, University of Osijek, Croatia
Nenad Šuvak, University of Osijek, Croatia

International Programme Committee

Austria: Peter Filzmoser, Vienna University of Technology
Croatia: Ksenija Dumičić, University of Zagreb
Nenad Šuvak, University of Osijek
Hungary: David Simon, Eötvös Loránd University Budapest
Italy: Paola Berchiolla, University of Torino
Slovenia: Luka Kronegger, University of Ljubljana

Published by: School of Mathematics and Informatics
University of Osijek

Edited by: Danijel Grahovac
Nenad Šuvak

Logo Design by: Andy Posner
L^AT_EX template: Vladimir Batagelj

Programme

Programme of the 27th Young Statisticians Meeting

Friday, September 29, 2023

18:00 – 19:30 Registration at School of Mathematics and Informatics

Saturday, September 30, 2023

8:00 – 8:30 Registration at School of Mathematics and Informatics

8:30 – 9:00 Opening of the Conference

9:00 – 10:30 **Session 1:** Statistics in Finance and Energy Markets

Chair: Ksenija Dumičić

- Ivana Cunjak Mataković (Croatia): *Selected statistical methods for outliers' detection for discovering financial manipulations with crypto-assets*
- Dominik Mihalčić (Croatia): *Models for natural gas consumption*
- Anja Žavbi Kunaver (Slovenia): *The challenge of multiple testing: Case of emission coupons trading*

10:30 – 11:00 **Coffee break**

11:00 – 12:30 **Session 2:** Statistics in Social Sciences

Chair: David Simon and Luka Kronegger

- Vili Smolič (Slovenia): *Linguistic analysis of suicide related questions in the online counselling service This is me*
- Michaela Bersia (Italy): *A Bayesian framework assessing temporal trends of suicidality among hospitalized adolescents during COVID-19 pandemic*
- Bendegúz Zaboretzky (Hungary): *Depression in times of a pandemic – the impact of COVID-19 on the lay discourses of e-mental health communities*

12:30 – 13:30 **Lunch**

13:30 – 14:30 **Session 3:** Mathematical Statistics

Chair: Peter Filzmoser

- Daniel Strenger (Austria): *Measuring dependence between a scalar response and a functional covariate*
- Patrick Benjamin Langthaler (Austria): *Quantifying and estimating dependence via sensitivity of conditional distributions*

14:30 – 15:00 **Coffee break**

15:00 – 16:00 **Session 4:** Mathematical Statistics

Chair: Nenad Šuvak

- Andrea Vlahek Štrok (Croatia): *Asymptotic behavior of an approximate maximum likelihood estimator of drift parameters in a multidimensional diffusion model*
- Michael Suja (Hungary): *Asymptotic theorems for contaminated runs of heads in the coin tossing experiment*

18:00 - 20:00 **Guided city tour**

20:00 **Conference dinner** (Hotel Osijek, Šamačka 3, 31000 Osijek)

Sunday, October 1, 2023

10:00 – 11:30 **Session 5:** Statistical methodology and medical statistics

Chair: Paola Berchiolla

- Maria Vittoria Chiaruttini (Italy): *Sample size calculation algorithm in two-stage three-arm sequential non-inferiority clinical trials: A simulation study*
- Piercesare Grimaldi (Italy): *Study of Risk Factors for Pulmonary Exacerbations in Patients with Cystic Fibrosis and the Effects of New CFTR Modulator Therapy on Chronic Bronchial Infection and Pulmonary Exacerbations in a Real-Life Study*
- Stefan Embacher (Austria): *Modelling antibody kinetics – When and how often to measure?*

11:30 **Closing of the conference and refreshment**

Abstracts

Selected statistical methods for outliers' detection for discovering financial manipulations with crypto-assets

Author: Ivana Cunjak Mataković¹

¹ *Centar-revizija d.o.o.*

Corresponding Author: ivana.cunjak@gmail.com

The first cryptocurrency ever was Bitcoin, which was described in a white paper in 2008: Bitcoin: A Peer-to-Peer Electronic Cash System, edited by Satoshi Nakamoto. From 2008 to the present, there are approximately 10,000 crypto assets worth about \$1.15 trillion. It seems that crypto-assets have become part of the financial culture.

The EU is one of the first which adopted comprehensive regulation of crypto-asset markets in May 2023. Article 91 of Regulation (EU) 2023/1114 of the European Parliament and of the Council on markets in crypto-assets is of particular interest: Prohibition of market manipulations, where manipulation occurs when misleading or false signals are given or the price of one or more crypto-assets is maintained at an abnormal or artificial level.

How can we detect market manipulations with crypto-assets? One possible way is to discover outliers. Detecting outliers has been a topic in statistics for centuries. In the last two decades, interest in methods for detecting outliers has increased particularly in the data mining community. Thus, there are different approaches to outliers in the literature depending on the research field and area.

In this research, Benford's Law will be used, which is based on the expected distributions of the first digits of the observed numbers. Benford's Law is one of the best-known preliminary methods for detecting collective outliers in the field of forensics. Research on the application of Benford's Law in the field of crypto-assets is ongoing. In this research, two of the most common market manipulation techniques with crypto-assets, "pump and dump" and "wash trading", are presented. The market manipulations with crypto-assets discovered by Benford's Law will be analysed. The research results can be used as a preliminary approach to detect suspicious financial transactions in order to protect investors, prevent the misuse of crypto-assets, and maintain market integrity.

Models for natural gas consumption

Author: Dominik Mihalčić¹

¹ *School of Applied Mathematics and Informatics, J. J. Strossmayer University of Osijek*

Corresponding Author: dmihalci@mathos.hr

Natural gas is primarily used for heating and industrial purposes. Predicting its consumption is of vital importance in gas distribution. It is mostly influenced by weather conditions (mainly temperature), but there is also a human factor which is hard to quantify. An inaccurate prediction, whether higher or lower than the actual consumption, results in unwanted expense for the gas distributor. Therefore, there is a need to create accurate forecasting models for gas consumption.

We primarily focused on Zagreb, whose consumption ranges from 1 to 25 million kWh. It is clear that tomorrow's consumption depends on today's and perhaps, to a lesser extent, on yesterday's consumption, so ARIMA model with regression was imposed as a logical candidate for a model. However, extremely large jumps in consumption occur during the colder period (sometimes up to 50%) so we had to look for a different approach. Partial improvement was provided using neural networks. We present several different models created along the way and discuss their strengths and weaknesses.

The challenge of multiple testing: Case of emission coupons trading

Author: Anja Žavbi Kunaver¹

Co-authors: Irena Ograjenšek²; Marjan Cugmas²

¹ *University of Ljubljana & GEN-I, d.o.o.*

² *University of Ljubljana*

Corresponding Author: anja.zavbi@gmail.com

In practice, we often face the challenge of multiple testing, which, without appropriate corrections, may lead to too many rejected true null hypothesis (type I error rate). The challenge also needs to be dealt with in the energy sector of the economy. One such case is trading with emission coupons, which we address in this study. The purpose of emission coupons is to reduce the emissions of carbon dioxide. Trading with coupons takes place at two levels: on the primary and secondary market. Here, companies which actively trade on the emission markets, strive to provide themselves with enough allowance for carbon dioxide emissions.

On the primary market, there are auctions every working day at 11 o'clock. The secondary market is active ten hours a day every working day. There is trading going on continuously on the secondary market and the traders want to know which day and which hour are optimal for selling or buying emission coupons (when is the actual price higher or lower than daily or weekly average).

There are various possible approaches for testing statistical significance of differences between average prices such as t-test, Wilcoxon sig-rank test, Tukey test and permutation test. In this study, we focus on comparison between paired t-test and a permutation test.

Our simulations show that both tests are appropriate and have about the same test power. But the paired t-test is faster and easier to use, so we choose it as more appropriate.

We also attempt to find out which multiple testing procedure is the most appropriate – with which procedure do we get the best test power and how big are differences between procedures.

Finally, we move away from simulations to real-life data in order to propose a relevant trading strategy.

Linguistic analysis of suicide related questions in the online counselling service This is me

Author: Vili Smolič¹

Co-authors: Marjan Cugmas ¹; Sara Atanasova ¹

¹ *Faculty of Social Sciences, University of Ljubljana*

Corresponding Author: vili.smolic@gmail.com

Social stigma and feelings of helplessness make it difficult for young people to seek face-to-face counselling for mental health problems and suicidality, however, due to its anonymity and convenience, the internet has proven to be an efficient tool in overcoming those barriers. Especially online health communities and online counselling services can be an important source of easily accessible information and anonymous support, including when it comes to mental health issues and problems. Unfortunately, due to a lack of face-to-face interaction and limited availability of non-verbal and visual cues, there is a decrease in mutual awareness of concerns and interests between patients and health professionals. To mitigate these limitations and optimize our effectiveness in providing patient care, it is beneficial to conduct a thorough analysis of posts within health community forums.

Therefore, the linguistic text analysis of more than 19 thousand questions and answers in the online counselling service will be presented. The data were obtained by the Slovenian largest and oldest online counselling service for young people This is Me for the period between 2012 and 2021. A special attention will be given to compare the sentiment of questions and answers among different sub-themes within the suicide questions, controlling for some users' characteristics such as gender and age. The analysis included word clouds, hierarchical clustering, sentiment analysis and multiple linear regression.

Results showed that themes related to problems at home and at school and themes related to serious mental illness emerge with suicide-themed questions. We found that counsellors' responses in general reflected a positive sentiment in suicide question types, although users had a highly more negative sentiment here. The sentiment in the service is also influenced by other factors such as gender and age.

Findings provide insight into the communication dynamics and counselling style in terms of sentiment between adolescents and counsellors, which could be helpful to the editors of such online counselling services.

A Bayesian framework assessing temporal trends of suicidality among hospitalized adolescents during COVID-19 pandemic

Author: Michela Bersia¹

Co-authors: Elisa Ferracin²; Emanuele Koumantakis¹; Lorena Charrier¹; Paola Berchiolla³; Paola Dalmaso¹; Rosanna Irene Comoretto¹

¹ *Department of Public Health and Pediatrics, University of Torino*

² *Regional Epidemiology Unit*

³ *Department of Clinical and Biological Sciences, University of Torino*

Corresponding Author: michela.bersia@unito.it

Introduction

Literature on temporal patterns of suicidality during the COVID-19 pandemic among youths is growing. The present work proposes a Bayesian approach to assess the COVID-19 pandemic temporal patterns of suicidality among inpatient adolescents.

Methods

Data referred to the first hospital discharge record with a suicidality-related ICD9-CM code among adolescents aged 13-19 between 1 January 2017 and 31 March 2021 were collected in the Piedmont region, Italy (n=334; median age: 15 years, IQR: 14-16; 80% girls). A Poisson Bayesian regression model performed on pre-COVID-19 data (2017-2020), adjusted by seasonality and stratified by gender, was adopted to provide a posterior predictive probability distribution, namely the distribution of the discharge counts per year quarter from April 2020 to March 2021 implied by the model. Then, using the posterior prediction on the generated model, the predictive posterior p-value (PPP) was assessed, defined as the probability that the observed number of cases in the COVID-19 period was equal to or lower than that expected according to previous years' distribution (or, conversely, the probability that the expected number of cases was higher than that observed in the COVID-19 period). We followed the guideline provided by Bland et al. 2015 to interpret PPP as a measure of evidence against the null hypothesis (equivalence between observed and predicted cases) (Bland, 2015).

Results

Among boys, we consistently obtained PPPs close to 1 throughout the entire 2020, suggesting that the observed phenomenon of suicidality was lower than what was expected. From January to March 2021, there was a high consistency between

the expected and observed phenomenon, as indicated by $PPP = 0.33$. In contrast, among girls, the PPPs, consistent with pre-pandemic data, had a sharp decrease in January-March 2021 ($PPP = 0.04$), suggesting strong evidence of a higher occurrence of the phenomenon than initially expected.

Conclusions

In 2020, a declining trend of suicidality was observed in April-June 2020 among both girls and boys, prolonged throughout the whole year among boys. Later, an increasing pattern of suicidality was registered in early 2021 (January-March) compared to the pre-pandemic period, more pronounced among girls than boys. The present findings encourage wider adoption of Bayesian approaches as valuable tools to explore rare events with low counts to enlighten open public health issues deeply.

Bibliography

1. Bland M. 2015. An Introduction to Medical Statistics (section 9.4). Oxford, UK: Oxford Univ. Press. 4th edition.

Depression in times of a pandemic – the impact of COVID-19 on the lay discourses of e-mental health communities

Author: Bendegúz Zaboretzky¹

Co-authors: Domonkos Sik ²; Eszter Katona ²; Renata Nemeth ²

¹ *Faculty of Social Science, Eötvös Loránd University*

² *ELTE Eötvös Loránd University*

The article aims at exploring the impact of the COVID pandemic on the lay discourses of depression emerging in online mental health forums. The narrative framing of depression plays a central role not only because it affects the instrumental strategies of the depressed subjects (e.g. preferred therapy), but also because it is a constitutive element of the identity of depressed subjects, thus affects the process of recovery itself. COVID had a serious impact on people living with mental disorders (especially depression and anxiety), thus our research aimed at mapping the consequences of these transformations on discursive level. A textual dataset of the most popular, English language online health forums was collected (n=339,550 publicly available entries posted between 15 February 2016 and 31 December 2020). Structural topic modelling was used to explore the various discursive patterns characterising the pre-and post-COVID era. Our results show that the pandemic did not take over the discursive space of depression forums, yet it transformed many aspects of it: a new horizon of critique opened up; the biomedical authority was reinforced; the ego-centric perspectives were refined; the previously unquestionable discursive frames become fragmented; and the horizon of emergency overshadowed the previous risk perspective.

Measuring dependence between a scalar response and a functional covariate

Author: Daniel Strenger¹

Co-author: Siegfried Hörmann¹

¹ *Graz University of Technology*

Corresponding Author: daniel.strenger@tugraz.at

Our aim is to extend the scope of a recently introduced dependence coefficient between scalar responses and multivariate covariates to the case of functional covariates. While formally the extension is straight forward, the limiting behavior of the sample version of the coefficient is delicate. It crucially depends on the nearest-neighbor structure of the covariate sample. Essentially, one needs an upper bound for the maximal number of points which share the same nearest neighbor. While a deterministic bound exists for multivariate data, this is no longer the case in infinite dimensional spaces. To our surprise, very little seems to be known about properties of the nearest neighbor graph in a high-dimensional or even functional random sample, and hence our main contribution is to advise a way how to overcome this problem. An important application of our theoretical results is a test for independence between scalar responses and functional covariates.

Quantifying and estimating dependence via sensitivity of conditional distributions

Author: Patrick Benjamin Langthaler¹

Co-authors: Jonathan Ansari ; Sebastian Fuchs ; Wolfgang Trutschnig

¹ *Paris-Lodron University Salzburg, Department of Artificial Intelligence and Human Interfaces*

Corresponding Author: patrickbenjamin.langthaler@stud.plus.ac.at

One of the oldest and classic statistical problems is the comparison of two groups by taking the difference of some characteristic of their underlying distributions. In this work we introduce a method for deriving a measure of dependence between two continuous random variables, X and Y , by taking the expected value of such a difference over all pairwise distributions of Y conditional on X . This method can therefore be used to derive measures of dependence between continuous variables from measures of pairwise differences between two distributions. We first show that several classical and more recent measures of dependence can be seen as special cases of this approach. We then show under what circumstances our novel method coincides with already existing methods. Finally we show how to construct new measures of dependence using our procedure. We show under which conditions these are true measures of dependence in the sense that they take values between 0 and 1, with 0 being attained only under independence and 1 being attained only under complete dependence, i.e. when Y is a measurable function of X . We provide a copula based approach for estimating these measures and prove the consistency of the estimator. Finally we present a simulation study to get an idea of the rate of convergence and the variance of the estimator, as well as a real-life data example for which our method can be used.

Asymptotic behavior of an approximate maximum likelihood estimator of drift parameters in a multidimensional diffusion model

Author: Andreja Vlahek Štrok¹

Co-authors: Miljenko Huzak²; Snježana Lubura Strunjak²

¹ Faculty of Chemical Engineering and Technology, University of Zagreb

² Faculty of Science, Department of Mathematics, University of Zagreb

Corresponding Author: avlahek@fkit.hr

For a fixed T and $k \geq 2$, we analyze a k -dimensional vector stochastic differential equation over the time interval $[0, T]$:

$$dX_t = \mu(X_t, \theta) dt + \nu(X_t) dW_t,$$

where $\mu(X_t, \theta)$ is a k -dimensional vector and $\nu(X_t)$ is a $k \times k$ -dimensional matrix, both consisting of sufficiently smooth functions. $(W_t, t \geq 0)$ is a k -dimensional standard Brownian motion whose components $W_t^1, W_t^2, \dots, W_t^k$ are independent scalar Brownian motions. Vector of drift parameters θ is unknown and its dependence is in general nonlinear. We prove that approximate maximum likelihood estimator of drift parameters $\hat{\theta}_n$ obtained from discrete observations $(X_{i\Delta_n}, 0 \leq i \leq n)$, when $\Delta_n = T/n$ tends to zero, is locally asymptotic mixed normal with covariance matrix that depends on maximum likelihood estimator $\hat{\theta}_T$ obtained from continuous observations $(X_t, 0 \leq t \leq T)$, and on path $(X_t, 0 \leq t \leq T)$. To prove the desired result, we emphasize the importance of the so-called uniform ellipticity condition of diffusion matrix $S(x) = \nu(x)\nu^T(x)$.

Similar considerations can be made for any sufficiently large $T > 0$. The main assumptions are then that X is an ergodic diffusion with stationary distribution and the diffusion matrix is diagonalizable so that only eigenvalues depend on X , i.e., there exists an orthogonal constant matrix U such that $S(x) = U\Lambda(x)U^{-1}$.

Asymptotic theorems for contaminated runs of heads in the coin tossing experiment

Author: Michael Suja¹

Co-author: István Fazekas¹

¹ *University of Debrecen*

Corresponding Author: michael.suja@science.unideb.hu

The classical coin tossing experiment is studied. Asymptotic theorems are obtained concerning the head-runs containing certain numbers of tails. It is proven that the limit distribution of the number of those runs of length n containing at most T tails is compound Poisson. The normalized first hitting time for the at most T contaminated head runs is shown to have an exponential limiting distribution just like any other waiting time distribution. However, accompanying distributions are obtained for the length of the longest head-runs containing at most T tails given that it does not have a limiting distribution. To this end, a two parameter family of accompanying distribution is offered. Simulation results supporting the theorems are also presented.

Sample size calculation algorithm in two-stage three-arm sequential non-inferiority clinical trials: A simulation study

Author: Maria Vittoria Chiaruttini¹

¹ *Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, Padova, Italy*

Corresponding Author: mariavittoria.chiaruttini@ubep.unipd.it

Non-inferiority studies play a crucial role in evaluating new treatments that offer potential benefits, such as reduced side effects, lower costs, or faster treatment times, compared to existing effective treatments. To determine whether a new treatment is not significantly worse than the reference treatment by more than a specified margin, non-inferiority studies are designed. However, a common criticism of such studies comparing experimental treatments with active controls is that the margin is often derived from existing literature rather than estimated based on the sample being studied. By including a placebo group in addition to the groups receiving the new treatment and the active treatment, the gold-standard design for non-inferiority studies circumvents this problem.

This research poses the interest on sample size calculation for a group-sequential design in a framework described above. In particular, based on the sample size estimation algorithm developed by Lanyu Lei (2020), that assumes normal outcomes and homoscedasticity across groups, the paper aims to i) assess the robustness of the sample size calculation algorithm in presence of deviation from normality, relaxing the homoscedasticity assumption, ii) introduce the stopping rule for futility at interim analysis, based on conditional power, and iii) describe the impact of the historical borrowing, through normalized power prior, for the placebo arm on Power and Type I error rate, both in case of “Smaller is better” and “Larger is better” design.

When the outcome variables are derived from three gamma distributions with equal variability, the statistical power consistently reaches an acceptable level using sample size calculated by the original algorithm. When $\sigma(P) > \sigma(A)$, using the sample size by the original algorithm, the study results overpowered, otherwise, introducing heteroskedasticity, the algorithm estimates a lower sample size without loss in power and probability to reject at interim. When $\sigma(A) > \sigma(P)$, using the sample size by the original algorithm the study results underpowered, otherwise, the proposed algorithm is able to keep stable the power and the probability to reject at interim. When the stop for futility is included, the adjustment of sample size

by the Inflation Factor (IF) can contrast the loss of overall power at the end of the enrolment. Historical borrowing introduces a bias, the larger the difference between the averages in the placebo arms the more marked the bias. In general, if the historical mean is lower than current average in placebo arm leads to decrease/increase in statistical power and type I error, on the other hand, when the historical mean surpasses current average in placebo arm leads to increase/decrease in statistical power and type I error, for the “Larger is better” and “Smaller is better” designs, respectively.

The study’s web-applications, RESKOUT and FUST-WING tools, provide a user-friendly interface for sample size and power calculations to help during the three-arm non-inferiority clinical trial designs.

Study of Risk Factors for Pulmonary Exacerbations in Patients with Cystic Fibrosis and the Effects of New CFTR Modulator Therapy on Chronic Bronchial Infection and Pulmonary Exacerbations in a Real-Life Study

Author: Piercesare Grimaldi¹

Co-authors: Paola Berchiolla²; Virginia de Rose³

¹ *Department of Public Health and Pediatrics, University of Torino*

² *Department of Clinical and Biological Sciences, University of Torino*

³ *Department of Molecular Biotechnology and Health Sciences, University of Torino*

Corresponding Author: piercesare.grimaldi@unito.it

Pulmonary exacerbations in cystic fibrosis (CF) patients have significant impacts on morbidity, mortality, and quality of life. This study aims to identify risk factors for exacerbations, evaluate the effects of new CFTR modulator therapy, and analyze factors influencing patient response to the therapy. This retrospective, single-center observational study involved CF patients treated with Ivacaftor (n=12) and untreated eligible patients (n=8). The impact of Ivacaftor on severe exacerbations and exacerbation incidence was analyzed using logistic regression and count models. The Zero-Inflated Model was used to account for excess zeros in the count data. Among patients treated with Ivacaftor, one experienced a severe exacerbation (9.1%), while none occurred in the untreated group. Logistic regression indicated that Ivacaftor treatment did not significantly affect severe exacerbation risk (p=0.621). The incidence rate ratio for exacerbations in treated versus untreated patients was 0.91 (p=0.897), suggesting no significant difference. Various predictors were analyzed, including genotypes, diagnosis history, treatment methods, and patient characteristics. Some predictors had significant effects on exacerbation rates, while others showed no significant associations. This study investigated the effects of Ivacaftor treatment on pulmonary exacerbations in CF patients. While some predictors showed associations with exacerbation rates, the overall impact of Ivacaftor treatment and other factors on exacerbation risk varied. These findings contribute to understanding the complex relationship between CFTR modulator therapy and exacerbation risk, helping guide personalized treatment approaches for CF patients.

Modelling antibody kinetics – When and how often to measure?

Author: Stefan Embacher¹

¹ *Medical University of Graz*

Corresponding Author: stefan.embacher@medunigraz.at

Infectious diseases play a substantial role in global public health concerns, and the evaluation of interventions aimed at reducing the burden of infection, such as vaccination, need rigorous studies to assess their effectiveness and efficiency. Measuring antibodies is a possibility to quantify the impact of a vaccination, and the modelling of antibody kinetics, i.e. the change of antibody titer over time, is crucial in optimizing the design of immunization trials. This includes determining the required sample size and optimal time points for antibody measurements/sampling in order to describe their kinetic patterns accurately. We are currently conducting a systematic review, with the main objectives of identifying studies that use mathematical or statistical models to describe antibody kinetics and to determine in which context and how they have been used in the process of study design. A range of mathematical and statistical models are employed to account for the typically longitudinal structure of the data in immunization trials. These models include fundamental statistical approaches, such as linear models, or more sophisticated statistical techniques like nonlinear mixed effects models. Mathematical models are also frequently used, mostly in the form of deterministic within-host compartmental models, with varying levels of complexity. Despite the relatively high number of included publications (270), hardly any considered models for study design.

In our research we try to answer the question of when and how often to measure antibody titer, to optimize the estimation of parameters used to model antibody kinetics. The sampling frequency depends on the number of parameters in the respective model, while promising approaches to determine the optimal sampling schedule utilize characteristics of the Fischer information matrix. For example, one of the most commonly used optimality criteria, D-optimality, finds the sampling schedule, which maximizes the determinant of the Fischer information matrix.

The limited use of mathematical models describing antibody kinetics, especially regarding study design, highlights the need and importance of basic research. Through our work, we aim to provide a framework, which can be actively used in practice to improve infectious disease study design.

Participants

1. **Vladimir Batagelj**, *vladimir.batagelj@fmf.uni-lj.si*, IMFM, University of Ljubljana, Slovenia
2. **Mirta Benšić**, *mirta@mathos.hr*, School of Applied Mathematics and Informatics, J.J. Strossmayer University of Osijek, Croatia
3. **Paola Berchiolla**, *paola.berchiolla@unito.it*, Department of Clinical and Biological Sciences, University of Torino, Italy
4. **Michela Bersia**, *michela.bercia@unito.it*, Department of Public Health, University of Torino, Italy
5. **Maria Vittoria Chiaruttini**, *mariavittoria.chiaruttini@ubep.unipd.it*, Department of Cardio-Thoraco-Vascular Sciences, University of Padova, Italy
6. **Rosanna Comoretto**, *rosannaire.comoretto@unito.it*, Department of Public Health and Pediatrics, University of Torino, Italy
7. **Ivana Cunjak Mataković**, *ivana.cunjak@gmail.com*, Centar-revizija d.o.o., Croatia
8. **Biserka Draščić Ban**, *biserka.drascic.ban@uniri.hr*, Faculty of Maritime Studies, University of Rijeka, Croatia
9. **Ksenija Dumičić**, *kdumicic.prof@gmail.com*, Faculty of Economics & Business, University of Zagreb, Croatia
10. **Srdan Dumičić**, *srdjan.dumicic@ipsos.com*, Ipsos West Adria, Croatia
11. **Stefan Embacher**, *stefan.embacher@medunigraz.at*, Medical University of Graz, Austria
12. **Anuška Ferligoj**, *anuska.ferligoj@fdv.uni-lj.si*, Faculty of Social Sciences, University of Ljubljana, Slovenia
13. **Peter Filzmoser**, *peter.filzmoser@tuwien.ac.at*, Institute of Statistics and Mathematical Methods in Economics Vienna University of Technology, Austria
14. **Herwig Friedl**, *hfriedl@tugraz.at*, Institute of Statistics, Graz University of Technology, Austria
15. **Danijel Grahovac**, *dgrahova@mathos.hr*, School of Applied Mathematics and Informatics, J.J. Strossmayer University of Osijek, Croatia
16. **Piercesare Grimaldi**, *pcgrimaldi@gmail.com*, Department of Public Health and Pediatrics, University of Torino, Italy
17. **Tina Košuta**, *tina.kosuta@mf.uni-lj.si*, Faculty of Medicine, University of Ljubljana, Slovenia
18. **Luka Kronegger**, *luka.kronegger@fdv.uni-lj.si*, Faculty of Social Sciences, University of Ljubljana, Slovenia
19. **Patrick Benjamin Langthaler**, *patrickbenjamin.langthaler@stud.plus.ac.at*, Department of Artificial Intelligence and Human Interfaces, Paris-Lodron University Salzburg, Austria
20. **Nyanga Honda Masasila**, *hondanyanga@gmail.com*, University of Debrecen, Hungary
21. **Hrvoje Mataković**, *hrvoje.matakovic@iztg.hr*, Institute for Tourism, Zagreb, Croatia

22. **Dominik Mihalčić**, *dmihalci@mathos.hr*, School of Applied Mathematics and Informatics, J.J. Strossmayer University of Osijek, Croatia
23. **Anna Molnár**, *mranna1817@gmail.com*, Eötvös Loránd University Budapest, Hungary
24. **Maximilian Ofner**, *m.ofner@tugraz.at*, Graz University of Technology, Austria
25. **Ivan Papić**, *ipapic@mathos.hr*, School of Applied Mathematics and Informatics, J.J. Strossmayer University of Osijek, Croatia
26. **Tamas Rudas**, *trudas@elte.hu*, Faculty of Social Sciences, Eötvös Loránd University Budapest, Hungary
27. **Kristian Sabo**, *ksabo@mathos.hr*, School of Applied Mathematics and Informatics, J.J. Strossmayer University of Osijek, Croatia
28. **Dávid Simon**, *simon.david@tatk.elte.hu*, Faculty of Social Sciences, Eötvös Loránd University Budapest, Hungary
29. **Vili Smolič**, *vili.smolic@gmail.com*, Faculty of Social Sciences, University of Ljubljana, Slovenia
30. **Janez Stare**, *janez.stare@mf.uni-lj.si*, Faculty of Medicine, University of Ljubljana, Slovenia
31. **Daniel Strenger**, *daniel.strenger@tugraz.at*, Graz University of Technology, Austria
32. **Michael Suja**, *michael.suja@science.unideb.hu*, University of Debrecen, Hungary
33. **Nenad Šuvak**, *nsuvak@mathos.hr*, School of Applied Mathematics and Informatics, J.J. Strossmayer University of Osijek, Croatia
34. **Andreja Vlahek Štrok**, *avlahek@fkit.unizg.hr*, Faculty of Chemical Engineering and Technology, University of Zagreb, Croatia
35. **Bendeguz Zaboretzky**, *zaboretzkybendeguz@gmail.com*, Eötvös Loránd University Budapest and visiting researcher at RC2S2 research group, Hungary
36. **Anja Žavbi Kunaver**, *anja.zavbi@gmail.com*, University of Ljubljana & GEN-I d.o.o, Slovenia

Supporters



Ipsos



Croatian Bureau of Statistics



Osijek-baranja County



City of Osijek



Tourist Board of the City of Osijek



**School of Applied Mathematics and Informatics
University of Osijek**



Foundation of the Croatian Academy of Sciences and Arts