BIOSTAT 2021

25th International Scientific Symposium on Biometrics

POREČ, CROATIA, **8 - 10 SEPTEMBER 2021**

HRVATSKO BIOMETRIJSKO DRUŠTVO - ZAGREB -

BOOK OF ABSTRACTS

EDITED BY: Anamarija Jazbec
Marija Pecina
Zdenko Sonicki
Diana Šimić
Mislav Vedriš
Slavica Sović

BIOSTAT 2021
25th ANNIVERSARY

ISSN 1849-434X

CROATIAN
established 1989. BIO
METRIC
SOCIETY

HRVATSKO
BIO osnovano 1989.
METRIJSKO
DRUŠTVO

BOOK OF ABSTRACTS

BIOSTAT 2021

25th International Scientific Symposium on Biometrics

POREČ, CROATIA, 8 -10 SEPTEMBER 2021

EDITORS:

Anamarija Jazbec
Marija Pecina
Zdenko Sonicki
Diana Šimić
Mislav Vedriš
Slavica Sović

Zagreb, 2021.

IMPRESSUM

BIOSTAT 2021
25th International Scientific Symposium on Biometrics
Poreč, Croatia, 8 - 10 September 2021

ORGANISER

Croatian Biometric Society

CO-ORGANISER

University of Zagreb
Faculty of Organization and Informatics
Varaždin

INTERNATIONAL SCIENTIFIC PROGRAM COMMITTEE

Anamarija Jazbec (Croatia)
Marija Pecina (Croatia)
Franjo Prot (Croatia)
Zdenko Sonicki (Croatia)
Diana Šimić (Croatia)
Bojana Dalbelo Bašić (Croatia)
Silvije Vuletić (Croatia)
Lynne Billard (USA)
Karina Gibert (Spain)
Annie Morin (France)
Mislav Vedriš (Croatia)
Johann Sölkner (Austria)
Janez Stare (Slovenia)
John E. Vena (USA)
Katarina Košmelj (Slovenia)
Nancy Flournoy (USA)

ORGANIZING COMMITTEE & EDITORIAL BOARD

Anamarija Jazbec, President of the OC
University of Zagreb, Faculty of Forestry
Marija Pecina
University of Zagreb, Faculty of Agriculture
Zdenko Sonicki
University of Zagreb, School of Medicine,
School of Public Health „Andrija Štampar“
Diana Šimić
University of Zagreb,
Faculty of Organization and Informatics
Mislav Vedriš
University of Zagreb, Faculty of Forestry
Slavica Sović
University of Zagreb, School of Medicine,
School of Public Health „Andrija Štampar“

23rd SCHOOL OF BIOMETRICS

Diana Šimić
Dept of Quantitative Methods
Faculty of Organization and Informatics
University of Zagreb

KEYNOTE SPEAKERS

Zvonimir Poljak
Ontario Veterinary College
University of Guelph

Karen L. Edwards
Dept of Epidemiology and Biostatistics
- Program in Public Health
Susan and Henry Samueli College of Health Sciences
University of California, Irvine

PUBLISHER:
Croatian Biometric Society
c/o University of Zagreb, School of Medicine,
School of Public Health „Andrija Štampar“
Rockefellerova 4, HR-10000 Zagreb, Croatia
www.hbmd.hr

VENUE:

Valamar Diamant Hotel
Brulo 1,
52440 Poreč, Croatia

FOR THE PUBLISHER:
Professor Zdenko Sonicki, PhD, President
University of Zagreb, School of Medicine,
School of Public Health „Andrija Štampar“
Rockefellerova 4, HR-10000 Zagreb, Croatia
hbmd@hbmd.hr

ISSN 1849-434X

BIOSTAT 2021
25th ANNIVERSARY

FULL REFERENCE:
Jazbec A, Pecina M, Sonicki Z, Šimić D, Vedriš M, Sović S, eds. (2021) Book of Abstracts BIOSTAT 2021 – 25th International Scientific Symposium on Biometrics, Poreč, Croatia, 8 – 10 September 2021. Zagreb: Croatian Biometric Society. ISSN 1849-434X

CONTENTS

BIOSTAT 2021
25th International Scientific Symposium on Biometrics
Poreč, Croatia, 8 - 10 September 2021

ABSTRACTS

PREFACE

BIOSTAT 2021

25th International Scientific Symposium on Biometrics
Poreč, Croatia, 8 - 10 September 2021

Welcome to the Book of Abstracts of BIOSTAT 2021 – The 25th International Scientific Symposium on Biometrics organized by the Croatian Biometric Society and co-organized by the University of Zagreb Faculty of Organization and Informatics Varaždin.

This silver jubilee, the twenty-fifth anniversary of Biostat, is going to be very specific, not just because it is twenty-fifth Biostat. It is going to take place in Poreč, Croatia, September 8-10, 2021, during COVID-19 pandemic crisis. In these challenging circumstances, known also as New Normal, with a lot of uncertainty, Biostat will be for the very first time organized as hybrid meeting. This means that, depending on the pandemic situation, all program sessions will be held both in person (in beautiful Poreč) and remotely (online - through the Zoom platform).

The symposium is an intrinsically interdisciplinary meeting accepting papers from a wide array of research disciplines at the intersection of data sciences and life sciences as evidenced by this Book of Abstract. BIOSTAT 2021 comprises 23rd School of Biometrics, keynote lectures, and contributed papers. This brochure contains 23+3 (7 students) abstracts written by 60 authors from 13 countries –Bosnia and Hercegovina, Canada, Czech Republic, Croatia, Germany, Italy, Norway, Slovak Republic, Slovenia, Spain, Switzerland, United Kingdom and United States of America.

The Symposium takes place on the Istrian peninsula in Poreč – a city with a thousand-year history.
Poreč developed on a small symmetrical peninsula long before the arrival of the Histrians, an Illyrian tribe. Today the centre of the old city is located on this small peninsula. With the arrival of the Histrians the area surrounding the present-day city was settled, and thanks to a naturally protected harbour the construction of a harbour-centre was enabled and Poreč could develop unhindered. Ever since the 1970s, the coast of Poreč and neighboring Rovinj has been the most visited tourist destination in Croatia.
We are continuing this year with our traditional 23rd School of Biometrics. Our guest lecturer this year is professor Diana Šimić Department of Quantitative Methods, Faculty of Organization and Informatics, University of Zagreb, with the topic of “Using Shiny Apps in Teaching Statistics”.

This year keynote speakers are professor Karen L. Edwards (Dept. of Epidemiology and Biostatistics - Program in Public Health Susan and Henry Samueli College of Health Sciences University of California, Irvine, USA) with a lecture on “Polygenic Risk Scores and prediction of metabolic syndrome: Is more better?”, and professor Zvonimir Poljak (Ontario Veterinary College, University of Guelph, Canada) with a lecture on “Predictive Analytics in Animal and Related Populations”.

Also, this year we are going on with awards for the best presentation for Students and for Scientists and encouraging BS and MS students to participate in Symposium with or without abstracts (presentations) with a popular registration fee of 0 € (0 kn).

We are looking forward to sharing with you pleasant moments, lively discussions, and successful networking, and to forging new partnerships and friendships. We believe that our exchange of ideas and interchange of knowledge will contribute to the progress of Biometric science.

Editors:
Anamarija Jazbec, Marija Pecina, Zdenko Sonicki, Diana Šimić, Mislav Vedriš and Slavica Sović
Zagreb, September, 2021

PROGRAM

BIOSTAT 2021

25th International Scientific Symposium on Biometrics
Poreč, Croatia, 8 - 10 September 2021

SYMPOSIUM PROGRAM at Glance

September, 7	19:00 – 20:00	Registration & Welcome Cocktail	
Wednesday, September, 8	9:00 – 9:30	Opening ceremony	Zdenko Sonicki, HBMD President Anamarija Jazbec, BIOSTAT OC President
	9:30 – 11:00	*Chair:*	Contributed session
	11:00 – 11:30	COFFEE BREAK	
	11:30 – 13:00	*Chair:*	Contributed session
	13:00 – 14:00	LUNCH BREAK	
	14:00 - 15:30	*Chair:*	School of Biometrics – *Part I* **Diana Šimić**: Using Shiny Apps in Teaching Statistics
	15:30 – 15:45	COFFEE BREAK	
	15:45 – 17:15	*Chair:*	School of Biometrics – *Part II* **Diana Šimić**: Using Shiny Apps in Teaching Statistics
	18:00	Antistress Fitness activities	
Thursday, September, 9	9:30 – 11:00	*Chair:*	Contributed session
	11:00 – 11:30	COFFEE BREAK	
	11:30 – 13:00	*Chair:*	Contributed session
	13:00 – 14:00	LUNCH BREAK	
	14:00 – 14:50	*Chair:*	Keynote Lecture: **Zvonimir Poljak** Predictive Analytics in Animal and Related Populations
	14:50 – 15:05	COFFEE BREAK	
	15:05 – 15:30	ISPC Meeting	
	16:00 – ??:??	OrganoLeptic and Audio Kinetic activities	
Friday, September, 10	12:00 – 13:00	LUNCH BREAK	
	13:00 – 14:20	*Chair:*	STUDENTS SESSION
	14:30 – 15:30	*Chair:*	STUDENTS SESSION
	15:30 – 16:00	COFFEE BREAK	
	16:00 – 17:00	*Chair:*	Contributed session
	17:00 – 17:50	*Chair:*	Keynote Lecture: **Karen L. Edwards** Polygenic Risk Scores and prediction of metabolic syndrome: Is more better?
	17:50 – 18:00	Closing ceremony	
	18:00 – 18:30	Farewell COFFEE	

Follow us at www.hbmd.hr

PROGRAM

BIOSTAT 2021
25th International Scientific Symposium on Biometrics
Poreč, Croatia, 8 - 10 September 2021

SYMPOSIUM PROGRAM in Detail

Tuesday, Sept, 7	19:00	Registration & Welcome Cocktail	
	20:00		
Wednesday, September, 8	9:00 – 9:30	Opening ceremony	Zdenko Sonicki, HBMD President Anamarija Jazbec, BIOSTAT OC President
	9:30 – 11:00	*Chair: Karina Gibert, Diana Šimić*	Contributed session
		AUTHOR	TITLE
		Želimir Kurtanjek	Statistical causality SCM in medicine
		Janez Stare	On goodness-of-fit of the Cox model
		Tomáš Mrkvička	What can be proven by functional test statistics: The R package GET
		Ana Perišić	A hybrid dissimilarity measure for mixed-type data clustering
	11:00 – 11:30	COFFEE BREAK	
	11:30 – 12:30	*Chair: Janez Stare, Mislav Vedriš*	Contributed session
		AUTHOR	TITLE
		Maja Buhin Pandur, Jasminka Dobša, Slobodan Beliga, Ana Meštrović	Topic modelling and sentiment analysis of COVID-19 related news on Croatian Internet portal
		Karina Gibert, Xavier Angerri	Evaluating the Impact of the COVID19 in Mental Health in people with Social Vulnerability through INSESS-COVID19 technology
		Andrija Štajduhar, Tomislav Lipić, Mirjana Kujundžić Tiljak	Understanding COVID-19 Pandemic in Croatia using Veridical Data Science
	12:30 – 14:00	LUNCH BREAK	
	14:00 – 15:30	*Chair: Zdenko Sonicki, Slavica Sović*	School of Biometrics – *Part I*: Diana Šimić: Using Shiny Apps in Teaching Statistics
	15:30 – 15:45	COFFEE BREAK	
	15:45 – 17:15	*Chair: Zdenko Sonicki, Slavica Sović*	School of Biometrics – *Part II*: Diana Šimić: Using Shiny Apps in Teaching Statistics
	18:00	Free activities	
Thursday, September, 9	9:30 – 11:00	*Chair: Marija Pecina, Anamarija Jazbec*	Contributed session
		AUTHOR	TITLE
		Gloria Gheno	Effects of temperature and other climatological variables on Covid19 infections
		Vladimir Brajković, Ivan Pocrnić, Miroslav Kapš, Marija Špehar, Maria Lie Selle, Vlatka Čubrić-Čurik, Strahil Ristov, Ivica Međugorac, Dinko Novosel, Gregor Gorjanc, Ino Čurik	Impact of the mitochondrial genome on quantitative traits in cattle
		Vostry L., Vostra-Vydrova H, Hofmanova B., Moravcikova N., Brajković V., Čubrić-Čurik V., Kasarda R., Čurik I.	Genome-wide association study of insect bite hypersensitivity in Old Kladruber horse

		Marija Špehar, Gerson A Oliveira Junior, Leticia A C Lara, Ivan Pocrnić, Christian Edel, Reiner Emmerling, Kay-Uwe Götz, Ino Čurik, Gregor Gorjanc	Dissection of genomic covariation of complex traits in dual-purpose Simmental cattle
	11:00 – 11:30	COFFEE BREAK	
	11:30 – 12:30	*Chair: Anamarija Jazbec, Mislav Vedriš*	Contributed session
		AUTHOR	TITLE
		Đodan Martina, Perić Sanja	Is selection of pedunculate oak provenances crucial for artificial regeneration of forest stands?
		Maja Moro, Mihaela Vuzem, Andreja Pirc Barčić, Darko Motik	Influence of the covid-19 pandemic on the employment trend in the croatian wood industry
		Azra Tafro, Luka Jurjević, Ivan Balenović	Validation of Allometric Remote Sensing Based Models for Pedunculate Oak Forests
	12:30 – 14:00	LUNCH BREAK	
	14:00 – 14:50	*Chair: Marija Pecina, Anamarija Jazbec*	Keynote Lecture: **Zvonimir Poljak** Predictive Analytics in Animal and Related Populations
	14:50 – 15:05	COFFEE BREAK	
	15:05 – 15:30	ISPC Meeting	
	16:00	Excursion	
Friday, September, 10	12:00 – 13:00	LUNCH BREAK	
	13:00 – 14:20	*Chair: Anamarija Jazbec, Marija Pecina*	Contributed session STUDENTS SESSION
		AUTHOR	TITLE
		Tea Sonicki, Markus Holzner	Generation of Tmem41b and ATG5 knockouts in mouse embryonic stem cell line
		Roko Bjelica	The effect of different soft tissues thickness of denture bearing area and different loading forces on peri-implant stress and strains in four mini-implant retained mandibular overdenture
		Sara Stalman, Zdenko Sonicki	Bloom Index – a measure of 'influencers' among scientists
		Lovro Marinović	Mortality differences between men and women after acute myocardial infarction
	14:30 – 15:30	Ljubica Vavan, Borut Bosančić	Biometrical approach in methodology of seed germination analysis in horticultural sciences modelled on *Gentiana lutea* L.
		Marta Horvat	Primary dysmenorrhoea among female students at the University of Zagreb – results and difficulties in conducting the survey
		Andrea Barbarić	Analysis of the influence of the observed variables on heart disease using logistic regression
	15:30 – 16:00	COFFEE BREAK	
	16:00 – 17:00	*Chair: Diana Šimić, Mislav Vedriš*	Contributed session
		AUTHOR	TITLE
		Meštrović T, Đaković I, Sonicki Z, Premužić Meštrović I, Pavić P	Exploratory data mining and retrospective studies in surgical practice – overcoming bias and other obstacles
		Nancy Flournoy	Advances in Adaptive Designs
	17:00 – 17:50	*Chair: Zdenko Sonicki, Slavica Sović*	Keynote Lecture: **Karen L. Edwards** Polygenic Risk Scores and prediction of metabolic syndrome: Is more better?
	17:50 – 18:00	Closing ceremony	
	18:00 – 18:30	Farewell COFFEE	

 # Note: Time for each title is 20 minutes - 15 for presentation followed by 5 minutes for questions and discussion

Meeting and networking: https://www.wonder.me/r?id=eef96458-81f7-454d-87ba-33c779c157a4

Follow us at www.hbmd.hr

23rd SCHOOL OF BIOMETRICS

23rd SCHOOL OF BIOMETRICS

Diana Šimić

Department of Quantitative Methods
Faculty of Organization and Informatics
University of Zagreb

Email: Diana.Simic@foi.unizg.hr
https://www.foi.unizg.hr/en/staff/diana.simic

Diana Šimić is a full professor of statistics and research methods in information sciences at the Faculty of Organization and Informatics University of Zagreb, and the Chair of the Postgraduate doctoral study program Information Sciences. She also teaches statistics at the postgraduate doctoral study program of Biomedicine at the School of Medicine University of Zagreb. She participated in 23 research projects, supervised four doctoral candidates; (co)authored more than 60 research papers, 11 books, and four book chapters. She was an invited lecturer at more than 20 conferences. She is a member of the Croatian Biometric Society, Croatian Statistical Association, Croatian Information Technology Association, Croatian Society for Medical Informatics, American Statistical Association (ASA), and Association for Computing Machinery (ACM). She serves as the vice president on the Boards of the Croatian Biometric Society and the Croatian Information Technology Association.

Topic:

Using Shiny Apps in Teaching Statistics

Abstract:

If picture is worth a thousand words, what's the worth of a moving picture? When teaching statistics, it is often easier to explain statistical concepts using simulation and visualization. Luckily, you can find many free interactive apps illustrating statistical concepts and procedures on the Internet. Still, there are many occasions when you would like to create your own app. Maybe your students do not speak English very well, or you would like to use your own data or example in the simulation, or you would just like to tweak one little aspect of the app to better suit your purpose. If you know basics of R and RStudio, and are familiar with ggplot2, you can easily create your own apps, which you can share with your students as files to be run in RStudio, or upload to a remote server and make accessible as web apps.

In this School of Biometrics, we will cover the basics of the rmarkdown and shiny packages, and show how to create and share your own stat app.

To participate actively in the school, you will have to install R and RStudio on your laptop. You will also have to install some R packages. All the school materials, including instructions, slides, examples, and assignments will be available at the github repository https://github.com/disimic/biostat2021.

KEYNOTE LECTURE

KEYNOTE SPEAKER

Zvonimir Poljak

Ontario Veterinary College
University of Guelph

Email: zpoljak@uoguelph.ca
https://ovc.uoguelph.ca/population-medicine/faculty/Zvonimir-Poljak

Topic:

Predictive Analytics in Animal and Related Populations

Abstract:

Over the last several decades, the way food animals are kept has considerably changed. The number of farms decreased, with simultaneous increase in the population size. Accompanied with general improvement in biosecurity; and in swine, separation of age groups to minimize frequency of contacts; these changes led to many health benefits. Some of them include increase in production efficiency and many aspects of food safety. However, frequent movement of animals between locations contributed, at least in part, to easier transmission of viral pathogens among distinct herds. Changes in herd demographics and in population size, and frequency of viral infections directly and indirectly contribute to the organization of veterinary services. A recent survey determined that among the viral diseases, porcine reproductive and respiratory syndrome virus (PRRSV), influenza A virus (IAV), and some porcine coronaviruses are among the top pathogens of concerns for veterinarians who practice swine medicine in North America. For many of these pathogens, partial genome sequence of the virus is an important information which contributes to decision making during management of clinical outbreaks. Such data could be used to attribute the outbreak source, and to forecast clinical impact of the pathogen in a population. Ultimately, this could inform design of infection and disease control practices. In our research group, we investigate approaches which could be used to predict clinical impact of specific viral genomes (e.g., PRRSV and IAV) at the population level, and assess the between-species transmission potential for influenza A viruses, including their zoonotic potential. In this presentation, I will be discussing some of the successes and challenges in investigating such questions using data-driven predictive modelling approaches.

References:

Keay S, Sargeant JM, O'Connor A, Friendship R, O'Sullivan T, Poljak Z. Veterinarian barriers to knowledge translation (KT) within the context of swine infectious disease research: an international survey of swine veterinarians. BMC Vet Res. 2020 Nov 2;16(1):416.

Melmer DJ, O'Sullivan TL, Greer A, Moser L, Ojkic D, Friendship R, Novosel D, Poljak Z. The impact of porcine reproductive and respiratory syndrome virus (PRRSV) genotypes, established on the basis of ORF-5 nucleotide sequences, on three production parameters in Ontario sow farms. Prev Vet Med. 2021 Apr;189.

KEYNOTE SPEAKER

Karen L. Edwards

Dept of Epidemiology and Biostatistics
- Program in Public Health
Susan and Henry Samueli College of Health Sciences
University of California, Irvine

Email: kedward1@uci.edu
https://www.faculty.uci.edu/profile.cfm?faculty_id=6052

Topic:

Polygenic Risk Scores and prediction of metabolic syndrome: Is more better?

Abstract:
Introduction and Methods: Susceptibility to most common chronic conditions is due to interactions between multiple genetic variants and nongenetic factors. Polygenic scores (PGS) are used to capture the complex genetic effects with the goal of identifying genetically susceptible individuals. PGS are often based on meta-analysis comprised of multiple genome-wide association analyses (GWAS) that emphasize common variants. However, there are many questions about how to best construct a PGS. For example, it is not clear how well PGS based on GWAS from European ancestry (EA) populations will predict disease in non EA populations. We use a multi-ethnic sample of families to: (1) evaluate the association between a common-variant genetic burden PGS and metabolic syndrome (MetS)-related traits; (2) to compare / contrast the prediction of MetS traits from PGS based under different scenarios. In this talk, I will introduce the concept of PGS and will utilize data from a large multi-ethnic family study of type 2 diabetes to illustrate several key points. Briefly, data from the GENNID study consists of 1502 subjects in 259 families from European-American (EA), Mexican-American (MA), African-American (AA), and Japanese-American (JA) families. The Clumping and Thresholding (C + T) method was used to construct PGS based on single ancestry-specific GWAS meta-analysis summaries for body mass index (BMI), summaries based on EA samples only, and from a meta-analysis consisting of multiple populations. We evaluated the association of PGS constructed under these different scenarios and using a linear mixed effects model (with fixed covariate effects and random family effects). P-values were approximated using a normal distribution. For each ethnic group, ROC curves and AUC values (using bootstrapping methods) were used to compare the ability of PGS to predict obesity (BMI>30).

Results and Conclusion: PGS based on a GWAS meta-analysis of multiple populations predicted better (i.e., had higher AUC) than PGS based on either ancestry-specific or EA GWAS. The AUC values ranged from 53% to 71%, but explained less than 6% of the variance in obesity. In general, the proportion of variance explained was low but is consistent with the literature for other common, complex conditions.

CONTRIBUTED SESSIONS

Title:

Statistical causality SCM in medicine

Authors:

Želimir Kurtanjek

Affiliation:

University of Zagreb, Faculty of Food Technology and Biotechnology*, Zagreb, Croatia

Abstract:

This contribution is a comparative study of statistical causal models CSM in view of large scale digitalization (big data) and networking in applied health and medical research. Modern medical diagnostics is profoundly subjected to computer supported knowledge and algorithmic statistical diagnostics. Especially important in research diagnostics is development of generalize artificial intelligence (GAI). Although causal effects are in focus of medical diagnostics, in medical practice they are rarely given explicitly and statistically validated. Development of mathematical theory of causality by J. Pearl is the major progress in basic etymology and algorithmic causality. It is based on a priori knowledge (hypothesis) and statistical evaluation of conditional probabilities of evidence data given as Bayes networks. Pearl proposition of hierarchy of causal inference follows three essentially stages: observation, intervention do(x), and counterfactually analysis. Presently AI algorithms (neural networks, decision trees) based on modelling of observation belong to the first stage of inference and due to numerous confounding data, like in medical diagnosis, are unable to yield causality and lead to medical intervention. To evaluate causality with AI models based only on data is confounded due to numerous, in medicine due to multiple risk factors, leading to mistake in intervention decisions. To eliminate confounding data has to be interpreted with directly acyclic graphs (DAG) of conditional probabilities as a structural causal model (SCM). Pearl introduced d-separation algorithm to manipulate joint probability function by which confounding of back door associations is eliminated. Here is applied methodology of structural causality algorithm analysis to analysis of longitudinal study of onset of diabetes mellitus available by The John Hopkins University School of Medicine, Diabetes Epidemiology Section. DAG model is evaluated by PC algorithm and validated by prediction of observed marginal probability functions. Applied is SEM structural equation model to predict linear causality coefficients. Determined are adjustment sets to close spurious back door associations. For each adjustment set by OLS determined are causal coefficients and compared to SEM predictions. Nonlinear interactions are modelled by categorical stratification and presumed multidimensional Gaussian distributions. The results are compared with partial dependence plots derived by gradient boosted rand forest model.

J. Pearl, D. Mackenzie, „The book of why”, Penguin Random House, UK, 2019.

*retired

Title:	**On goodness-of-fit of the Cox model**
Authors:	**Janez Stare**
Affiliation:	University of Ljubljana, Faculty of Medicine
Abstract:	A standard instruction for those using the Cox model to analyze survival data is: After having fitted a Cox model one should consider to check proportional hazards and the functional form. In this presentation I will argue that these two assumptions are not independent and I will stress that the proportional hazards assumption is conditional on the functional form being correct and that testing for the functional form makes no sense if the ph assumption is satisfied.

Title:

What can be proven by functional test statistics: The R package GET

Authors:

Tomáš Mrkvička

Affiliation:

University of South Bohemia

Abstract:

Statistical testing is one of the major tools in biostatistics. Usually, the test statistic is one-dimensional, gathering the information from time or space into a single number. Performing the statistical inference with functional test statistics (such as the slope of warming measured every day of the year, spatial correlation measured in certain distances, F statistic of the GLM measured in every voxel of the brain) can reveal more information than the single agglomerative test statistic. On the other hand, using functional test statistics brings difficulties in the test statistic model assumptions, such as normality or homogeneity. Therefore, we have introduced a powerful, nonparametric statistical inference method with functional test statistics in our R package GET. The methods also provide graphical inference, which is equivalent to the formal inference, which allows for easy interpretation of the results. The package provides inference for functional GLM with one-, two- or three-dimensional functions, goodness-of-fit test based on multiple functional test statistics, graphical comparison of several distribution functions, graphical functional clustering, graphical test of dependence of two variables, functional central region detection together with functional box plot. It also allows for composite hypothesis testing in goodness of fit testing, i.e., when the model parameters must be estimated. All the procedures satisfy the family wise error rate control. We are recently working on false discovery rate control to detect all hypothesis (domain of functional test statistic) that should be rejected. The provided procedures are based on the ordering of functions according to extreme rank length functional depth, which allows for intrinsic graphical interpretation. The intrinsic graphical interpretation means that if the functional test statistic lies in at least one point outside the constructed envelopes, the null hypothesis is rejected. Thereafter, it identifies the domain of rejection.

- Myllymäki M., Mrkvička T. (2019). GET: Global envelopes in R. http://arxiv.org/abs/1911.06583
- Myllymäki M., Mrkvička T., Seijo H., Grabarnik P., Hahn U. (2017). Global envelope tests for spatial processes, JRSS Series B 79/2, 381-404.
- Mrkvička T., Roskovec T., Rost M. (2021). A nonparametric graphical tests of significance in functional GLM, Methodol Comput Appl Probab. 23, 593-612.

Title:	**A hybrid dissimilarity measure for mixed-type data clustering**
Authors:	**Ana Perišić**
Affiliation:	Polytechnic of Šibenik
Abstract:	One of the greatest challenges in clustering mixed-type data is finding the adequate distance function between objects. Most distance metrics work either with continuous-only or categorical-only data, but in applications, however, mixed-type data are prevalent in many real-world applications. Hybrid distance methods involve selecting a distance function that can accommodate mixed-type variables where a popular hybrid distance function is Gower's distance. This work presents a hybrid dissimilarity measure for mixed-type data where distances are calculated conditional on the feature type. The proposed dissimilarity measure is established as a normalized linear combination of distances following the principles of Gower's coefficient calculation. For numerical features, distances are calculated by applying a modified winsorized Huber loss, while for categorical features, a distance measure based on variable entropy is incorporated. The established measure is robust to outliers, skewed and sparse data, and can handle unbalanced categorical features and highly skewed numerical features.

Title:	**Topic modelling and sentiment analysis of COVID-19 related news on Croatian Internet portal**
Authors:	**Maja Buhin Pandur[1], Jasminka Dobša[1], .Slobodan Beliga[2,3], Ana Meštrović[2,3]**
Affiliation:	[1]Faculty of Organization and Informatics, University of Zagreb [2]University of Rijeka, Department of Informatics [3]Center for Artificial Intelligence and Cybersecurity University of Rijeka, Croatia
Abstract:	The aim of research is to identify topics and sentiments related to COVID-19 pandemic in Croatian online news media. For analysis we used news related to COVID-19 pandemic from Croatian portal *Tportal.hr* published- in the period from 1st January 2020 till 19th February 2021. Topic modelling was conducted by usage of the LDA method. Dominant emotions and sentiments related to extracted topics were identified by National Research Council Canada (NRC) sentiment emotion lexicon created originally for English, and translated into Croatian, among other languages. The results of this research enable better understanding of the crisis communication in the Croatian media related to the COVID-19 pandemic. This study is performed within the InfoCoV project funded by the Croatian Science Foundation.

Title:

Evaluating the Impact of the COVID19 in Mental Health in people with Social Vulnerability through INSESS-COVID19 technology

Authors:

Karina Gibert and Xavier Angerri

Affiliation:

Intelligent Data Science and Artificial Intelligence Research Center and Institut de Ciència i Tecnologia de la Sostenibilitat, Universitat Politècnica de Catalunya—BarcelonaTech, 08001 Barcelona, Spain;

Abstract:

The project INSESS-COVID19 aims to provide sufficient evidences on the impact of pandemics on social vulnerability to assess the Catalan government on mental health and other 19 areas of life. It develops a technological infrastructure to allow people from the entire territory to easily participate in the study and a sophisticated knowledge model on the data set metainformation to lead the intelligent data analysis and automatic reporting. It is providing information for strategical decisions, even in front of an unexpected systemic disruption, as COVID19 was. A technological cloud infrastructure and an instrument based on the SSM vulnerability model were developed for the public consultation. SSM includes 17 different faces of vulnerability like Mental Health. Our instrument had 9 variables specific for mental health. It was implemented in form of a digital questionnaire in a secure web. The web contained all additional materials to provide the citizen sufficient context of the consultation and instructions even to overcome digital gap issues. A sophisticated software was implemented to perform the automatic analysis and reporting. In 25 minutes the software analyses and creates a document with the results of the analysis and the diagnosis in the executive working document to be directly used in strategic meetings.

Consultation was spreadout by the 107 local areas of Social Services in Catalonia, each of them targeting a minimum of 20 citizens from July 2020 to December 6th 2020 About 1000 citizens responded with a 24% of people with mental health issues.

In December 15th 2020, only 9 days after, final results were publicly presented. Regarding mental health, 41.45% of participants required psychological support during COVID19. The loneliness feeling increased by 29.3% of participants and this is a pattern mainly followed by women (70%) older than 60 years (on average), mainly living alone, with some digital gap, from which a 52.18% required emotional support during the pandemics. On the other hand 49.64% of participants teleworked and from them, a 34.34% of cases also required emotional support. About 24% of the participants declared to have a diagnoses on mental health, a 96.94% of them are getting medication. From those, 46.28% declares depression, 58.51% anxiety, 10% severe mental health disorder, 6.1% Limit Personality disorder, 4.8% Post Traumatic disorder, 11.4% other mental disorders. These were the first quantitative evidences of the impact of pandemics on mental health in Catalonia.

The research had developed with advanced data science and AI tools, New indicators were created giving additional information. About 5.12% of participants debuted with some mental issues during the COVID19 first lock down. A 73.78% had mental health problems, 68.86% with depression, 72.38% with anxiety and 54% with Intellactual disability declare to feel worse by July 2020. Mental health was among the most relevant impacts produced by COVID19 crisis affecting all kind of people.

Title:

Understanding COVID-19 Pandemic in Croatia using Veridical Data Science

Authors:

Andrija Štajduhar[1], Tomislav Lipić[2], Mirjana Kujundžić Tiljak[1]

Affiliation:

[1]Andrija Štampar School of Public Health,
School of Medicine, University of Zagreb
[2]Ruđer Bošković Institute

Abstract:

The recent pandemic continues to plague our society, while nations are making tremendous efforts to respond effectively and mitigate risks through targeted actions, from the individual to the collective level. Because these actions have long-term impacts on society, they must be planned and executed with care. To support the global response to novel strains of coronavirus, various relevant data are being collected and shared, providing unique resources to scientific and data driven research communities and policy makers. Understanding and properly interpreting these data is the most important tool in the fight against the pandemic and its all-pervasive impact on our lives. Here, we use openly available governmental, socio-economic, political and public health data and investigate the temporal dynamics of the COVID-19 pandemic in Croatia with the aim of understanding, explaining, visualizing and predicting its nationwide and EU-wide impact under the Croatian policy regime. Motivated by the recently proposed Predictability, Computability and Stability (PCS) framework for Veridical Data Science, we present a new data analysis methodology that answers specific questions about the course of the pandemic in Croatia at the national and subnational levels. We build on and extend the principles of statistics, machine learning, and scientific inquiry, and embed the scientific principles of prediction and reproducibility into data-driven decision making. We estimate the causal impact of the application of health policies, social distancing measures, geographic mobility, and public events on outbreak spread metrics. We provide a comparison across EU countries, an overview of efforts by other approaches, and an assessment of the effectiveness of the methods used. Our work supports well the methodological foundations of Veridical Data Science and trustworthy AI needed to solve problems of great societal importance.

Title:

Effects of temperature and other climatological variables on Covid19 infections

Authors:

Gloria Gheno

Affiliation:

Innovative data analysis (Italy), Ronin Institute (USA)

Abstract:

The study of the factors which influence the spread of Covid19 has been addressed by many scholars in the last two years. A relationship between temperature and virus propagation was highlighted globally. Most studies agree on the existence of a negative association between Covid19 infections and ambient temperature. Only a small minority of works finds a non-monotonic or non-existent relationship or a positive association. An explanation for these discordant results can be attributed to the improper use of the average temperature which in states with a larger extension, such as the USA and China, is not a good indicator, because their territory includes areas with a significant climatic difference. The relationship between temperature and virus propagation is not sufficient to explain the differences in the spread of the virus among the various environmental zones of the planet. Of course, the different propagation of the virus is also attributable to the political differences of governments, to the various cultural, economic and population density differences. The climate, however, plays a fundamental role in the spread of Covid19, in Europe, for example, in the summer 2020 period the percentage of infections has decreased, increasing in autumn and winter. In this work I study how the climate affects the propagation of the coronavirus, analyzing different states of the 5 continents, divided into climatic zones if with large territorial extensions. I use De Martonne's table for the classification of their various climates. I carry out this analysis with methodologies for time series and with regression models, unlike the other studies, which do not use them together. In particular for generalized linear models (GLM), I propose a beta regression introducing a new response function which makes possible the analysis of non-monotonic relationships and in which monotonicity is determined not a priori but directly from the data. This flexibility helps to study and validate the widely declared relationship between temperature and the percentage of infections from Covid19. I estimate the beta regression parameters by maximizing the likelihood function with my modified version of the genetic algorithm and with the bootstrap method I get the standard errors. Knowing the influence of climatic factors on the spread of the virus globally can help the governments of the various states to make certain and adequate political choices, for example targeted vaccination and lockdown plans.

Title:

Impact of the mitochondrial genome on quantitative traits in cattle

Authors:

Vladimir Brajkovic[1], Ivan Pocrnic[2], Miroslav Kaps[1], Marija Špehar[3], Maria Lie Selle[4], Vlatka Cubric-Curik[1], Strahil Ristov[5], Ivica Medugorac[6], Dinko Novosel[1], Gregor Gorjanc[2], Ino Curik[1]

Affiliation:

[1]University of Zagreb Faculty of Agriculture, Svetosimunska 25, 10000 Zagreb, Croatia
[2]University of Edinburgh, The Roslin Institute, Easter Bush, Midlothian, Scotland, UK, EH25 9RG
[3]Centre for Livestock Breeding, Svetošimunska cesta 25, 10000 Zagreb, Croatia
[4]Norwegian University of Science and Technology, Faculty of Information Technology and Electrical Engineering, Alfred Getz' vei 1, 7034 Trondheim, Norway
[5]Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia
[6] LMU Munich, Faculty of Veterinary Medicine, Lena-Christ-Str. 48, D-82152 Martinsried, Germany

Abstract:

Intensive efforts have been made to improve production in livestock by implementing genomic selection based on autosomal markers. Here, we present the impact of mitochondrial genome on milk production traits of Holstein cattle, which is rarely studied. Complete mitogenomes representing 109 maternal lines were sequenced using NGS technology. This allowed the assignment of mitogenome sequence information to 2393 cows with 7115 milk production records (milk, fat and protein yield). A quantitative genetic model was applied to estimate the proportion of total variance explained by mitogenome inheritance (m_2) using five different models: (1) cytoplasmic model with maternal lineages (m_2CYTO), (2) haplotypic model with mitogenome haplotypes (m_2MITO), (3) amino acid model with unique amino acid combinations (m_2AMIN), (4) evolutionary model based on BEAST phylogenetic analysis (m_2EVOL) and mitogenome SNP model (m_2SNP). The effects of animal, sex chromosome, parity, calving season, region, year, heard and age at first calving were also included in each model. The estimated proportions of phenotypic variances explained by mitogenome inheritance (m_2CYTO, m_2MITO, m_2AMIN and m_2SNP) ranged from 4% to 7% for all three milk traits. At the same time, the estimated proportions of phenotypic variances explained by m_2EVOL and by sex chromosomal loci were negligible. The obtained results indicate that a considerable proportion of phenotypic variance in milk traits is explained by mitogenome variation. A better understanding of the use of mitogenome inheritance can play an important role in practical breeding and remains a challenge.

Keywords: milk production traits, mitochondrial genome, variance components, next generation sequencing (NGS), Holstein cattle

Title:	**Genome-wide association study of insect bite hypersensitivity in Old Kladruber horse**
Authors:	**Vostry L.[1], Vostra-Vydrova H.[1,2], Hofmanova B.[1], Moravcikova N.[3], Brajkovic V.[4], Cubric-Curik V.[4], Kasarda R.[3], Curik I.[4]**
Affiliation:	[1]Czech University of Life Sciences, Kamycka 129, 165 21 Prague, [2]Institute of Animal Science, Pratelstvi 815, 10400 Prague, Czech Republic, [3]Slovak University of Agriculture in Nitra, Tr. A. Hlinku 2, SK 949 76 Nitra, Slovak Republic, [4]University of Zagreb, Faculty of Agriculture, Svetošimunska cesta 25, 10000 Zagreb, Croatia
Abstract:	Insect bite hypersensitivity (IBH) is the most common allergic skin disease in horses, and the etiology is multifactorial. The occurrence of the disease is influenced by environmental factors and complex genetic inheritance. The initial disease stimulus is a bite by midges of *Culicoides spp*. or, less commonly, *Simulium spp*. Knowledge of associated genes is limited. The objective of this study was to identify genomic regions (genes) significantly associated with the occurrence of IBH (yes/no) in Old Grey Kladruber horses. A genome-wide association study (GWAS) was conducted using data from a pedigree and 220 genotyped horses (GGP Equine 70k chip). After SNP pruning, the final data set consisted of 65,157 autosomal loci. GWAS was performed using the linear mixed model approach based on a genomic relationship matrix calculated from SNP marker frequency. Fixed effects analysis was performed with a generalized linear model using a one-step genomic selection approach. This approach projects genomic relationships onto all ungenotyped individuals in a pedigree (constructing the so-called H-matrix). After correction for multiple comparisons, we identified nine regions on five different chromosomes (10, 11, 15, 20, and 26) significantly associated with the occurrence of IBH. IBH was largely inherited as a polygenic trait, with several regions showing significant effects, some of which had not been previously reported. The reported results will be used in the creation of a selection strategy for IBH status in the Old Kladruber Horse. The study was supported by projects QK1910156, MZE-RO0718, APVV-17-0060 and ANAGRAMS-IP-2018-01-8708 (some methodological parts).

Title:

Dissection of genomic covariation of complex traits in dual-purpose Simmental cattle

Authors:

Marija Spehar[1], Gerson A Oliveira Junior[2], Leticia A C Lara[3], Ivan Pocrnić[3], Christian Edel[4], Reiner Emmerling[4], Kay-Uwe Götz[4], Ino Curik[5], Gregor Gorjanc[3,6]

Affiliation:

[1]Croatian Agency for Agriculture and Food, Svetošimunska 25, Zagreb 10000, Croatia
[2]Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, ON, Canada
[3]The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland, UK
[4]Institute of Animal Breeding, Bavarian State Research Center for Agriculture, 85586 Grub, Germany
[5]University of Zagreb Faculty of Agriculture, Department of Livestock Sciences, Svetosimunska 25, Zagreb 10000, Croatia
[6]Animal Science Department, Biotechnical Faculty, University of Ljubljana, Groblje 3, 1230 Domzale, Slovenia

Abstract:

Work on multiple-selection is based on knowledge about covariance and correlation among traits. In this study, we used ~9K progeny tested dual-purpose Fleckvieh bulls genotyped with ~50K markers and analysed its dairy (milk yield - MY and fat yield - FY) and beef traits (net daily gain - NG and carcass grading - CG) in three steps. First, we estimated overall genetic parameters (covariances) between the dairy and beef traits at the whole-genome level. Second, we performed temporal analysis by estimating temporal changes in genetic covariances and correlations. Third, we partitioned covariances among dairy and beef traits by chromosome, chromosome regions and individual loci to genic covariance, within chromosome linkage disequilibrium (LD) and between chromosome LD. First, we estimated allele substitution effects of SNP markers with a multivariate marker model (ridge regression). For this step we used a Monte Carlo Markov Chain method and saved samples from posterior distribution of allele substitution effects. From these allele substitution effects we calculated breeding values from which we calculated covariance components which were basis for correlations calculation. Allele substitution effects were positively correlated between most trait combinations (rMY: FY =0.65, rMY: NG =0.07, rFY: NG =0.03, and rNG: CG =0.52), with an exception of rMY: CG =-0.10 and rFY: CG =-0.12. Overall correlations within dairy and beef traits were high and positive (rMY: FY =0.81, rNG: CG =0.52), while they were low to moderate between these two groups of traits (rMY: NG =0.23, rMY: CG =-0.05, rFY: NG =0.18 and rFY: CG =-0.10). A detailed analysis of the trends shows that most of the changes are driven by the build-up of a negative LD (the Bulmer effect) for individual traits as well as for pairs of traits, while corresponding genic covariances changed very little. In the future we can optimize mattings in such way to exploit covariation due to within and between chromosome LD.

Title:

Is selection of pedunculate oak provenances crucial for artificial regeneration of forest stands?

Authors:

Đodan Martina, Perić Sanja

Affiliation:

Croatian Forest Research Institute

Abstract:

Pedunculate oak (Quercus robur L.) is the most important tree species in Croatia from commercial viewpoint. Regeneration of p. oak forest stands is demanding with increasing number of challenges which add to the management complexity. The share of artificially regenerated oak stands is high (two thirds in seedlings production on the national scale) for which, selection of appropriate oak provenance could be important. If poorly chosen, direct consequences for forest management could result in low wood production, low vitality, financial losses and problems with regeneration. To address these challenges selection of appropriate p. oak provenance must meet not only current requirements of site conditions, but future conditions as well. In contrary, established forest stand will not have needed stability to reach the end of rotation. Thus, provenance tests provide necessary background for silvicultural decision making in terms of provenances' growth, production, vitality and consequently nursery production plans. The aim of the research is to answer the question is there significant differences in tree height and diameter at breast height (DBH) between provenances and if so, to single out the best provenances which could be supported and the worse ones to be excluded from management. Research includes 16 provenances on the locality Gajno (trial plots in central Croatia). Provenances originate from natural distribution of p. oak in Croatia, from the most eastern edge (provenance Mitrovica, Serbia, just at border with Croatia) up to the most western distribution edge (provenance Motovun). Trial was established in four repetitions, with 16 different provenances in the spring of 1988. In the spring of 2021. all trees in the trial were measured (tree height with Haglöf/Vertex IV and DBH with calliper). Trees were carefully measured according to their unique number for the purposes of repeated measure in 2023 and calculation of height, diameter and volume increment. Data were analysed using Statistica soft. ANOVA showed significant differences among studied provenances. Data analysis revealed that selection of provenance must be based on scientific results since it has prominent role in growth and development of p. oak trees in the fourth decade of their development. These findings are strongly advised to be used for artificial regeneration of p. oak stands in Central Croatia as well as for planning of nursery production on the national scale.

Title:	**INFLUENCE OF THE COVID-19 PANDEMIC ON THE EMPLOYMENT TREND IN THE CROATIAN WOOD INDUSTRY**
Authors:	**Maja Moro, Mihaela Vuzem, Andreja Pirc Barčić, Darko Motik**
Affiliation:	Faculty of Forestry and Wood Technology, University of Zagreb, Croatia
Abstract:	This article presents an analysis of the current situation in the Croatian woodworking sector on the basis of established values of labour resources in Wood processing and Furniture manufacturing as well as anticipating the influence of COVID-19 on the employment trends. According to Croatian Standard Classification of Occupations (NKD 2007), for sectors Wood processing (C16) and Furniture manufacturing (C31), the data of the number and composition of persons in paid employment in legal entities are gathered from Croatian Bureau of Statistics. The data of the number of daily new COVID-19 cases are gathered from the official government website for accurate and verified information on Coronavirus in the Republic of Croatia. The aim of this paper is to find the answer to how the Covid-19 pandemic affected the movements of employment in the two main Croatian wood sectors.

Title:	**Validation of Allometric Remote Sensing Based Models for Pedunculate Oak Forests**
Authors:	**Azra Tafro[1], Luka Jurjević[2], Ivan Balenović[2]**
Affiliation:	[1] Faculty of Forestry and Wood Technology, University of Zagreb, Croatia [2] Croatian Forest Research Institute, Jastrebarsko/Zagreb, Croatia
Abstract:	Light detection and ranging systems (LiDAR) that use laser light to measure distances are becoming increasingly popular in modeling and estimating forest attributes. By providing larger datasets than traditional field measurements they enable better model estimation, and modern computational methods allow for more complex models. However, there are currently no general guidelines for methods of testing and comparing the models. In this work, based on airborne LiDAR data and field measurements in lowland pedunculate oak forests of Pokupska basin, we provide a comprehensive overview of models and (cross-)validation methods and propose a possible best-practice standard in this setting.

Title:	**Exploratory data mining and retrospective studies in surgical practice – overcoming bias and other obstacles**
Authors:	**Meštrović T[1], Đaković Bacalja I[2], Sonicki Z[3], Premužić Meštrović I[4], Pavić P[1]**
Affiliation:	[1]University Hospital Center Zagreb – "Rebro", Croatia [2]General Hospital Bjelovar, Croatia [3]School of Public Health „Andrija Štampar“, Croatia [4]Clinical Hospital "Merkur", Zagreb, Croatia
Abstract:	Retrospective studies are still frequently used in surgical domains, due to their relative speed and low cost in comparison to randomised prospective clinical trials. They are, however, prone to different sorts of bias and are also very dependent on sample size and the quality of data, which are frequently missing in medical records. The data are usually analysed by classical regression methods that do not tolerate a small sample with a large number of predictors and missing values, while stepwise regression causes overfitting. It is in solving this problem that the value of data mining, especially "ensemble" methods, is manifested. In particular, random forest methods generate models of high predictive value, selecting important variables, detecting exceptions in a set of examples, and cleaning the data from errors. An approach of selecting variables for subsequent multivariate statistical analysis by data mining methods, that are relatively resistant to the abovementioned obstacles, was tested in two vascular surgical domains: for creating outcome prediction rules for ruptured abdominal aortic aneurysms (rAAA) and for unravelling risk factors for development of vascular surgical site infections (VSSI). While in the first of the abovementioned domains, we were able to create a relatively simple preoperative prediction model for mortality in rAAAs, which was prospectively validated; in the case of a study that searched for specific risk factors for vascular surgical site infections after surgical vascular reconstructions, we encountered problems in validation due to shift in bacterial landscape over time. The research will, therefore, be expanded according to the influx of new data with potentially relevant variables, as well as with the analysis of risk factors for the emergence of VSSIs with specific resistant and non-resistant bacteria. The subsequent analyses will also be performed in order to find the source of bias and to try to improve on the methodology for retrospective analysis of medical records in surgical domains, and to compensate for the less than ideal datasets that we frequently have to deal with.

Title:

Advances in Adaptive Designs

Authors:

Nancy Flournoy

Affiliation:

University of Missouri

Abstract:

Our understanding of adaptive designs has expanded in several directions recently. Similarities are emerging in the distribution theory of designs in which treatment allocation probabilities adapt and designs with sample size re-estimation and/or early stopping options. We will discuss the uses of observed versus Fisher information in analysis as exemplified in an enrichment study by Lin, Flournoy and Rosenberger that is in press at the Annals of Statistics which builds on prior work:

1. Flournoy, Nancy, Caterina May, and Chiara Tommasi. "The effects of adaptation on maximum likelihood inference for nonlinear models with normal errors." *Journal of Statistical Planning and Inference* 214 (2021): 139-150.
2. Lin, Zhantao, Nancy Flournoy, and William F. Rosenberger. "Random norming aids analysis of non-linear regression models with sequential informative dose selection." *Journal of Statistical Planning and Inference* 206 (2020): 29-42.

We will comment on our expanding understanding of how the observed information might be used more strategically in design, drawing from

1. A Lane, Adaptive designs for optimal observed Fisher information, *Journal of the Royal Statistical Society: Series B* 82: 1029-1058, 2020

We will connect important features of these works with our new theoretical understandings about the effects of sample size re-estimation and early stopping options, drawing from

1. Tarima, Sergey, and Nancy Flournoy. "Most Powerful Test Sequences with Early Stopping Options." *arXiv preprint arXiv:2101.11595* (2021).
2. Tarima, Sergey, and Nancy Flournoy. "Choosing Interim Sample Sizes in Group Sequential Designs." *German Medical Data Sciences: Bringing Data to Life*. IOS Press, 2021. 11-16.
3. Tarima, Sergey, and Nancy Flournoy. "Asymptotic properties of maximum likelihood estimators with sample size recalculation." *Statistical Papers* 60.2 (2019): 373-394.
4. Tarima, Sergey, and Nancy Flournoy. "Distribution theory following blinded and unblinded sample size re-estimation under parametric models." *Communications in Statistics-Simulation and Computation* (2019): 1-12.

STUDENT SESSION

Title:

Generation of Tmem41b and ATG5 knockouts in mouse embryonic stem cell line

Authors:

Tea Sonicki, Markus Holzner

Affiliation:

Institute of Molecular Health Sciences, Department of Biology, Swiss Federal Institute of Technology, ETH Hönggerberg, Zurich, Switzerland

Abstract:

Smoothened is a G-protein coupled receptor, a key component of the hedgehog signaling pathway, and plays a vital role in the embryonic development and adult tissue homeostasis. Dysfunctional signaling often leads to severe forms of medulloblastoma and other developmental issues. In a screen for SMO modulators Tmem41b came up as hit. Literature suggests that it is involved in the regulation of autophagy and lipid metabolism but the mechanism by which it regulates HH signaling is not clear yet. Elucidating it, a Tmem41b knockout cell line was generated, and functional assays were performed. To further determine and control for the role of autophagy in this mechanism, we additionally generated knockout ESCs for ATG5. Autophagy is an intracellular process degrading and recycling unwanted cytoplasmic proteins, macromolecules and organelles, and ATG5 is necessary for the initiation of this process.

In this study we generated Tmem41b and ATG5 knockouts in embryonic stem cells with the CRISPR/Cas9 system using guide RNAs targeting the introns 1 and 3 of both genes, resulting in the deletion of the second and third exon. To confirm derived clones as knockouts, 2 pairs of primers were used for both genes, with one pair targeting DNA outside of the deleted region while the other pair amplified the to be deleted region. In total, multiple knockout clones were generated for the ATG5 and Tmem41b genes. To validate these clones, protein and transcriptional expression were determined and functional tests performed.

Taken together, this project generated knock out ES cell lines for two genes and hereby bases future studies unraveling the mechanism by which Tmem41b executes its cellular functions.

Title:	**The effect of different soft tissues thickness of denture bearing area and different loading forces on peri-implant stress and strains in four mini-implant retained mandibular overdenture**
Authors:	**Roko Bjelica**
Affiliation:	School of Dental Medicine University of Zagreb
Abstract:	**Background and aim:** Soft tissues overlying edentulous residual ridges are supposed to be the primary stress bearing area, affecting the stability and amount of denture movements under masticatory loads. Dental implants are used to enhance removable denture retention and stability, especially in geriatric patients with thin residual ridges. The aim of this study was to evaluate the influence of different thickness of mucosal tissues under different mandibular overdenture loadings on peri-implant stress and strains of four mini-implants (MDI). **Methods:** A model of the mandible with severely atrophied and thin residual ridge was 3D printed, based on a CBCT scan, using a material with similar mechanical properties as the bone. Four MDIs (2.0x10 mm) were inserted interforaminaly at approximate sites of teeth 34, 32, 42, and 44, parallel to each other. Strain gauges (SG) were placed at the cervical MDI sites, one SG buccally and another orally. Altogether, eight SGs were placed. Measurements were made with artificial mucosa (soft silicone) of 1.5, 3 and 4 mm thickness overlying posterior denture bearing area. Forces of 50N, 100N and 150N were applied on the overdenture in the first molar region bilaterally. Each measurement was repeated three times. The effects of tissue thickness (1.5, 3, 4 mm), loads (50, 100, 150 N), and their interaction on peri-implant stress and strains was investigated using a multivariate factorial analysis of variance (MANOVA) and Scheffe post hoc test. Statistical program SPSS 20.0 was used for statistical analysis. **Results:** A combination of compressive and tensile deformations were recorded with SGs and the computer software (Kyowa DCS-100A Dynamic Data Acquisition Software) showing the maximum values for each of the applied forces. Multivariate factorial ANOVA revealed significant influence of tissue thickness, loads and their interaction ($p<0.001$). With an increase of the applied loads, an increase in deformations was recorded ($p < 0.001$). By increasing the artificial mucosa thickness from 1.5 to 3 mm, the peri-implant deformations increased, but with the 4 mm thick mucosa stresses further increased in some SG positions, but in other turned to negative values ($p < 0.001$). **Conclusions:** Within the limitations of this study, it was concluded that the thickness of artificial mucosal tissues had significant effect on strain and stresses of cervical peri-implant bone. Increased mucosal thickness and loads may cause adverse effects to both, peri-implant bone or MDIs. **Key words:** multivariate factorial ANOVA, mini-implants, overdenture, mucosal thickness

Title:

Bloom Index – a measure of 'influencers' among scientists

Authors:

Sara Stalman[1], Zdenko Sonicki, MD, PhD[1,2]

Affiliation:

[1]School of Medicine, University of Zagreb, Croatia
[2]Andrija Stampar School of Public Health

Abstract:

INTRODUCTION: Living in the era of social media, there is a risk that this form of communication is gaining too high a value as people are more prone to believe publicly exposed scientists than those ‘working in silence’. This study investigates discrepancy between certain scientist's appearance in the media and the number of citations they have, inspired by Neil Halls’s paper on Kardashian Index[1].

METHODS: We chose 15 scientists from the Scientific Council of the Government and searched for the number of followers on Twitter, Facebook and Instagram for each one. Facebook and Instagram were excluded from the research due to inactivity of chosen scientists. As Twitter is not that popular among Croats, we altered the ‘celebrity value’ (CV) of Kardashian index[1] by counting the amount of statements each scientist made in four most read Croatian online journals[2] including 24sata, Večernji list, Dnevnik.hr and Tportal. The data was being collected from 19th of April 2021 until 8th of May 2021. After adding up the numbers, we took it as a measure of modified CV and named it ‘influencer value’ (IV). Furthermore, we investigated the number of citations for each scientist found on the Web of Science, in the same period of time as mentioned previously, and took it as a measure of ‘scientific value’ (SV). Based on Hall’s Kardashian Index[1], we developed a surrogate version – Bloom Index.

RESULTS: The results show that most scientists in Croatia are not considered ‘Science Kardashians’ as both of their indices are low. 1 out of 15 Scientific Council’s members has a high Bloom Index, as expected considering his position. Results also show that scientists in Croatia are more prone to make a public appearance in the news media than on social networks.

CONCLUSION: We conclude that today’s media makes it very easy for scientists to build an impressive image of themselves by simply having an opinion on something and ‘being louder’ than the others. On the other hand, those with high-quality work remain hidden from the world and in times like these, where we are in constant need of accurate and truthful informations, they should, perhaps, get more involved in the media as to spread a word about their commendable work.

[1] https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0424-0
[2] https://rating.gemius.com/hr/tree/8 (based on the data from April 2021)

Title:

Mortality differences between men and women after acute myocardial infarction

Authors:

Lovro Marinović

Affiliation:

School of Medicine, University of Zagreb

Abstract:

Background: Acute myocardial infarction is one of the most common causes of death in developed countries. Previous studies have shown that women experience higher mortality in the period following acute myocardial infarction. It is still not clear whether this is due to higher average age of female patients or due to other causes.

Objective: To determine the association of female sex to mortality after acute myocardial infarction during 36-48 months of follow-up.

Patients and methods: This study used patient data from the MIRCO registry (Myocardial Infarction Registry of Comorbidities and Outcomes) of patients who were treated for acute myocardial infarction in University Hospital Centre Sestre milosrdnice. For this study we used patient data from the year 2015.

Results: Among 451 patients with acute myocardial infarction, 320 (71%) were male and 131 (29%) were female. The average age of male and female patients was 63 and 73, respectively. Male patients were more likely to smoke, be overweight and have hyperlipidemia. Female patients had higher average glucose blood levels on admission (all $p<0.05$). The survival rate during follow-up was 85.0% for male patients and 77.1% for female patients. Kaplan-Meier curves showed that female patients had significantly higher mortality than male patients. However, after using multivariate Cox regression to account for the higher average age of women, the difference became statistically insignificant (HR=1.50 [0.92-2.45], $p=0.104$).

Conclusion: Women have a lower incidence of acute myocardial infarction than men. However, in the following years they have a significantly higher mortality rate. This difference can be explained by the higher average age of female patients. To further study this difference, larger and multicentric studies should be done.

Keywords: acute myocardial infarction, sex differences, mortality

Title:	**Biometrical approach in methodology of seed germination analysis in horticultural sciences modelled on *Gentiana lutea* L.**
Authors:	**Ljubica Vavan[1,2], Borut Bosancic[1,2]**
Affiliation:	1. University of Banja Luka Faculty of Agriculture, 2. University of Banja Luka Genetic Resources Institute
Abstract:	Seed germination analysis is regular work for most horticultural species. It depends and varies through combination of different phonological, ecological and anthropological factors. It is especially difficult in species where seed is collected from wild populations, such as *Gentiana lutea* L. This requires specific approach which includes literature study and field work. There are no specific protocols to approach those situations. Especially in literature review where variability of data and approaches deems contradictory conclusions. In this study proposed is specific system of formalised literature review in form of meta-analysis. This is followed by experimental tests in laboratory and then the controlled field testing. Initial results indicated success when such approach is applied and lack of success when it was ignored. Even slight differences in effect size and significance in meta-analysis have proven valuable source of information for further laboratory and field tests.

Title:

Primary dysmenorrhoea among female students at the University of Zagreb – results and difficulties in conducting the survey

Authors:

Marta Horvat

Affiliation:

School of Medicine, University of Zagreb

Abstract:

Primary dysmenorrhoea is one of the most common menstrual problems affecting women of reproductive age. It has a huge impact on academic as well as work performance, social activities and relationships of affected women, but it still isn't usual to talk openly neither about menstruation nor about menstrual pain.

The aim of the conducted study was to determine the prevalence, impact and pain management of primary dysmenorrhoea among female students at the University of Zagreb.

At the beginning, a request for cooperation was sent to deans of all components at the University of Zagreb to get their consent on doing a survey among students at their faculty or academy. The response rate was only 64,7%. After getting the consent, a web-based questionnaire was used to collect the study data among students. A total of 1210 female students voluntarily filled out the questionnaire, but 92 have been excluded due to preset exclusion criteria and other conditions. The exclusion criteria were age above 30 years, due to which nine students were excluded, and former diagnosis of a pathological condition related to secondary dysmenorrhoea, which ejected another 44 students. Furthermore, 16 students were excluded because there was no consent for participation given from their faculty management and 23 more were excluded due to an incomplete or incorrectly filled out questionnaire. A total of 1118 students was included in study analysis.

The prevalence of primary dysmenorrhoea observed in this study was 91,4%. Almost all students with dysmenorrhoea, 97,3% of them reported that it limited their daily activities in some way. In terms of pain management, 87,4% of the students were using medications, while nonpharmacological methods of relieving pain were used by 76,5%.

Instead of presenting just results obtained in the study, this presentation will also be about some difficulties occuring during its conduction and data „lost“ due to lack of consent and exclusion of some participants.

Keywords: dysmenorrhoea, menstrual pain, university students, survey, exclusion criteria

Title:	**Analysis of the influence of the observed variables on heart disease using logistic regression**
Authors:	**Andrea Barbarić**
Affiliation:	University of Zagreb, Faculty of Science, Department of Mathematics - Graduate study of Mathematical Statistics
Abstract:	In this part of my graduate thesis, logistic regression was applied to concrete data. Specifically, the database consists of 297 observations representing patients and 14 variables representing information about their health status at hospital admission (available on https://www.kaggle.com/cherngs/heart-disease-cleveland-uci). Based on this, the aim was to assess which variables have statistically significant effect on the occurrence of heart disease. Univariate and multivariate logistic regression and Stepwise procedure were used in the development of the most adequate model. The results have shown that heart disease is mostly affected by gender, chest pain type, resting blood pressure, the slope of the peak exercise ST-segment, number of major vessels colored by fluoroscopy, thallium test results and exercise induced angina.

BIOSTAT 2021
25th ANNIVERSARY

CROATIAN
established 1989. BIO
METRIC
SOCIETY

HRVATSKO BIOMETRIJSKO DRUŠTVO · ZAGREB ·

HRVATSKO
BIO osnovano 1989.
METRIJSKO
DRUŠTVO