STRengthening Analytical Thinking for Observational Studies (STRATOS): Neutral comparison simulation studies as the cornerstone to compare statistical methods

Anne-Laure Boulesteix (1,2), Mark Baillie (3), Dominic Edelmann (4), Leonhard Held (5), Tim Morris (6), Willi Sauerbrei (7) on behalf of the Simulation Panel

(1) Institute for Medical Information Processing, Biometry and Epidemiology, Faculty of Medicine, Ludwig Maximilian University of Munich, Munich, Germany

(2) Munich Center of Machine Learning, Munich, Germany

(3) Novartis Pharma AG, Basel, Switzerland

(4) Division of Biostatistics, German Cancer Research Center, Heidelberg, Germany

(5) Department of Biostatistics at the Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland

(6) MRC Clinical Trials Unit at UCL, London, UK

(7) Institute of Medical Biometry and Statistics, Medical Center - University of Freiburg, Freiburg, Germany

In Issue 2/2020 the simulation panel of the STRATOS initiative was introduced. In the second paragraph it reads 'It is obvious that simulation studies, and the more complex concept of 'neutral comparison studies' (Boulesteix et al 2017), are and will remain a key instrument to systematically assess and/or compare competing statistical methods and to create solid evidence to support STRATOS guidance. The panel had published a letter to the Editors of Biometrical Journal (the journal of the IBS-GR, -ROeS and -IR) entitled "On the necessity and design of studies comparing statistical methods" (Boulesteix et al, 2018). The first goal of this letter was to point out the importance of neutral comparison studies, the second goal was to stress the necessity to study the methodology of such comparison studies, in particular the design and the assumptions underlying simulation studies.

To further promote this type of studies and the improvement of their methodology, a team gathering four STRATOS members (ALB, MB, TM, WS) and two further biostatisticians committed to the topic (DE, LH) edited a special collection that recently appeared in Biometrical Journal entitled "Towards neutral comparison studies in methodological research."

As outlined in the editorial (Boulesteix et al., 2024), *"[b]iomedical researchers are frequently faced with an array of methods they might potentially use for the analysis and/or design of studies. It can be difficult to understand the absolute and relative merits of candidate methods beyond one's own particular interests and expertise. Choosing a method can be difficult even in simple settings but an increase in the volume of data collected, computational power, and methods proposed in the literature makes the choice all the more difficult. In this context, it is crucial to provide researchers with evidence-supported guidance derived from appropriately designed studies comparing statistical methods in a neutral way, in particular through well-designed simulation studies.*

While neutral comparison studies are an essential cornerstone toward the improvement of this situation, a number of challenges remain with regard to their methodology and acceptance. Numerous difficulties arise when designing, conducting, and reporting neutral comparison studies. Practical experience is still scarce and literature on these issues almost inexistent. Furthermore, authors of neutral comparison studies are often faced with incomprehension from a large part of the scientific community, which is more interested in the development of "new" approaches and evaluates the importance of research primarily based on the novelty of the presented methods. Consequently, meaningful comparisons of competing approaches (especially reproducible studies including publicly available code and data) are rarely available and evidence-supported state of the art guidance is largely missing, often resulting in the use of suboptimal methods in practice.

In this context, the intention of this collection was to gather (and stimulate the production of) studies in the field of biometrics that aim to fill this gap and, as such, can be regarded as atypical. We called for the submission of

1. well-designed neutral comparison studies of methods (including but not limited to studies arising from community challenges), that is, comparison studies fulfilling the two following criteria: (i) focused on the comparison of existing methods already described elsewhere rather than on a new prototype method being introduced; (ii) authored by a group of researchers who are (ideally) approximately equally familiar with all the compared methods;

2. papers defining, developing, discussing, or illustrating concepts related to practical issues and improvement of neutral comparison studies in the context of methodological biometrical research, including but not limited to the design, analysis, and presentation of reliable simulation studies, study protocols, study registration and (structured) reporting, replication studies, uncertainty quantification, and research synthesis. Papers of this type will provide a lens through which to critically reflect on neutral comparison studies in the future."

The special collection includes as many as 23 high-quality contributions presenting neutral comparison studies on various biometrical topics as well as metascientific contributions addressing the methodology of such studies. A special session with three talks from authors of the special collection and a panel discussion was organized at CEN23 in Basel by Sarah Friedrich. The topic attracted a lot of attention and stimulated fruitful discussions at the conference. Our expectations were overall clearly exceeded!

The Simulation Panel contributed a paper of the second type ("papers defining, developing, discussing, or illustrating concepts") entitled "Phases of methodological research in biostatistics— Building the evidence base for new methods" (Heinze et al., 2024). In this framework, methodological studies are viewed as contrib-

uting to "evidence on methods" in a similar way as the different clinical trial phases in drug development. Importantly, the paper stresses the importance of late phase studies – consisting of the extensive study of methods in a neutral way in various settings with the aim to understand when it is recommended or not. We believe such methodological guidance deserves more attention and space in major (bio)statistical journals.

We hope the success of our special collection will be the beginning of a paradigm shift changing the way the scientific community as a whole addresses the process of method development and evaluation to generate more reliable empirical evidence.

References:

Boulesteix AL, Baillie M, Edelmann D, Held L, Morris TP, Sauerbrei W, 2024. Editorial for the special collection "Towards neutral comparison studies in methodological research." Biometrical Journal 66:2400031.

Boulesteix AL, Binder H, Abrahamowicz M, Sauerbrei W, 2018. On the necessity and design of studies comparing statistical methods. Biometrical Journal 60: 216-218.

Boulesteix AL, Wilson R, Hapfelmeier A, 2017. Towards evidence-based computational statistics: lessons from clinical research on the role and design of real-data benchmark studies. BMC Medical Research Methology 17:138.

Heinze G, Boulesteix AL, Kammer M, Morris TP, White IR, 2024. Phases of methodological research in biostatistics—Building the evidence base for new methods. Biometrical Journal 66: 2200222..

Region News

Argentinian (RArg)

Between 8-10 October 2024, the Argentine Group of Biostatistics (GAB), representing the Argentine Region of the IBS, will convene the XXVIII Scientific Meeting of the GAB. The event will be hosted at the Faculty of Natural Sciences, University of Buenos Aires. Distinguished national and international speakers, including Dr. Raúl Macchiavelli (University of Puerto Rico), Dr. Juan Antonio Carbonell-Asins (Institute for Health Research, Valencia, Spain), Dr. Florian Hartig and Dr. Maximilian Pichler (University of Regensburg, Germany), Dr. Pablo Inchausti (University of the Republic, Uruguay), Dr. Anabel Forte (University of Valencia, Spain), Andreas Mayr (University of Bonn, Germany), Elena Ieno (Highland Statistics Ltd.), Dr. Javier Mariani (Huésped Foundation, Argentina), Dra. María Llames (University of San Martín, Argentina), Dr. Walter Sosa Escudero (University of San Andrés, Argentina), MSc. Iván Barberá (University of Comahue, Argentina), Dr. Pablo Turjanski, Dr. Regino Cavia, Dra. Maria Eugenia Szretter Noste, Dra. Lucía Babino and MSc. Valeria Gogni (University of Buenos Aires, Argentina), among others, will enrich the event with their expertise. The Meeting agenda will feature keynote lectures, short courses covering specific topics, workshops, and the Young Biometricians Contest.

Currently, Argentina is facing significant socioeconomic obstacles, especially regarding the funding of science and technological development. Despite these limitations, we are firmly committed to strengthening the capacities and bonds of our scientific community dedicated to biostatistics and data science, convinced of its essential contribution to national development.

Silvia Sühring
Biometric Bulletin Correspondent

Australasian (AR)

IBS-AR Student Scholarships

To help attract enthusiastic and talented students to career paths in biometrics, the Australasian Region offers scholarships for suitably qualified students who intend to undertake a fourth or honours year of study, or a coursework Masters, in statistics, mathematical statistics, biostatistics, bioinformatics or biometrics.

This year we had several excellent applications and we are delighted to announce the winners: Peter Orlovskiy (Masters student at the University of Auckland, New Zealand) and Ryan Borges (Honours student at the University of Sydney). Congratulations!

Biography - Peter Orlovskiy

After strongly considering studying actuarial science or economics, I stumbled upon, and found myself in, data science; I found the challenges satisfying, the mental stimulation sometimes intense, and quickly became fascinated with how binary machines can accomplish tasks like word generation, language translation, image classification, and video segmentation. Since then, I have not looked back.

Currently, I'm studying for a Master of Data Science at the University of Auckland, where I'll be under Joerg Wicker and Katerina Taskova's excellent supervision. My research aims to use adversarial learning to explore existing fisheries models' applicability domains and create a methodology/criterion to estimate their future performance after encountering distributional shifts.

My current interests include adversarial attacks, biologically inspired spiking neural networks, and the conditions under which adding stochastic elements to deep networks can help improve robustness against adversarial attacks. However, I am also interested in reinforcement learning, knowledge representations, machine understanding, and reasoning. I regularly attend machine learning seminars and participate in discussions about new findings and thoroughly enjoy learning about SOTA techniques and methodologies.

Outside of data science, I love reading about black holes and dark matter, and being the nerd that I am, I love watching math videos on YouTube. Physically, I enjoy hiking, swimming, and climbing mountains. I play tennis and love my early morning saunas. I am an avid fan of science fiction and epic space fantasy, and my friends and