

STRENGTHENING ANALYTICAL THINKING FOR OBSERVATIONAL STUDIES - A VERY BRIEF UPDATE ON THE ACHIEVEMENTS OF THE STRATOS INITIATIVE IN THE LAST 5 YEARS

Willi Sauerbrei¹, Michal Abrahamowicz², Mark Baillie³, Bianca De Stavola⁴, Mitchell Gail⁵, Marianne Huebner⁶, Ruth Keogh⁷ and Pamela A. Shaw⁸ for the STRATOS initiative

(1) Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Germany (2) Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada (3) Biostatistical Sciences and Pharmacometrics, Novartis Pharma AG, Basel, Switzerland (4) GOS Institute of Child Health, University College London, U.K. (5) Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, USA (6) Department of Statistics and Probability, Michigan State University, East Lansing, USA (7) Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK (8) Division of Biostatistics, Kaiser Permanente Washington Health Research Institute, USA

It is 5 years since Sauerbrei, Huebner, Collins, Lee, Freedman, Gail, Goetghebeur, Rahnenfuehrer and Abrahamowicz introduced the STRATOS initiative in the *Biometric Bulletin* (BB, [2017, 34 \(3\)](#)) to members of IBS. Each co-author represented one of the 9 topic groups of STRATOS. Authors of the current very brief update are the members of the Executive Committee and the chairs of the Publication Panel.

The STRATOS initiative was launched in 2013 at the 44th meeting of the International Society for Clinical Biostatistics (ISCB) in Munich and the first STRATOS paper summarized the motivation, mission, structure and aims of this international initiative (Sauerbrei, Abrahamowicz, Altman, le Cessie and Carpenter, 2014, [www.http://stratos-initiative.org/](http://stratos-initiative.org/)). Providing accessible, evidence-based guidance for key topics in the design and analysis of observational studies is the main aim. Guidance is intended for applied statisticians and other data analysts with varying levels of statistical background and experience. The focus is on health sciences research, but the content is also relevant for applications of statistics in other empirical sciences.

STRATOS was influenced by reporting guidelines, for more than a decade coordinated by the Enhancing the QUALity and Transparency Of health Research (EQUATOR) network and is an intellectual child of ISCB (Sauerbrei, Abrahamowicz, le Cessie, 2016). Soon after STRATOS started its activities, the *Lancet* published the series of articles entitled 'Research: increasing value, reducing waste' which heavily criticized research in the health sciences, including issues related to design, conduct, analyses, incomplete and unusable reports,

all of them being highly relevant for our profession and society at large (Kleinert and Horton 2014). Doug Altman was one of the co-authors and, in fact, his editorial in the *British Medical Journal* 'The scandal of poor medical research – we need less research, better research, and research done for the right reason' had already raised many issues twenty years earlier (Altman 1994). This visionary leader, who died in June 2018, was also one of the driving forces of the STRATOS initiative. Fortunately, Doug had inspired many younger colleagues and students with his thoughts about the essential role of critical appraisal and improvements in the quality of methodological and medical research (Sauerbrei et al 2020). Many have followed in his steps, as recently illustrated in the paper by van Calster, Wynants, Riley, van Smeden and Collins (2021), four of whom are STRATOS members. These authors provided an overview of issues resulting from the current organization of science that lead to research waste. They name practices resulting from prioritizing publication acceptance over publication quality. They also provided examples of initiatives aiming to improve the methodology and reproducibility of research. The STRATOS initiative is one of the thirteen topics or initiatives listed, and the only one with a strong focus on methods of statistical analysis (van Calster et al 2021).

State of the art recommendations and guidance needed

Statistical methodology has seen substantial development in recent times. Unfortunately, many of these developments are often ignored in practice. Even worse, 'standard' analyses reported in the medical literature are often based on unrealistic assumptions or use unsuitable methods, casting doubt on their results and conclusions. Recommendations concerning state-of-the-art methodology and guidance are needed. To improve statistical methodology and its transparency, statistical researchers must put more emphasis on comparing competing strategies and must generate evidence to support state-of-the-art methodologies. They must also provide guidance that is appropriate for the large community of researchers who perform data analysis but with varying statistical knowledge and experience.

In 2013 STRATOS started up with seven topic groups (TGs) focusing on different aspects of study design and analysis methodology. For their specific topic, each group provided a brief summary of the state of research, main issues, main aims and planned future research (Sauerbrei et al 2014). Two further TGs were initiated in 2015 on the topics of Survival analysis (TG8) and High-dimensional data (TG9). Summaries are available on the STRATOS website. Several TGs also have their own websites, where more detailed information is provided, including links to other resources.

All TGs have tackled very difficult issues and have made strong progress. For example, TG3 (Initial data analysis) and TG1 (Missing data) have derived frameworks for their topics, TG4 (Measurement error and misclassification) has derived guidance for less and more complex methods for error adjustment. However, some TGs are still far from offering guidance. For example, the overview of TG2 (Selection of variables and functional forms in multivariable analysis) classified seven topics needing further research, and projects working on guidance for researchers with basic statistical knowledge have only started (see Guidance and education below). Since the first BB article in 2017, all TGs have summarized their work and plans in short BB articles. The following table provides an overview of BB articles and links them to our website. Members of TGs and panels are listed in the articles.

TG1: Missing Data	https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#stratos Carpenter J, Lee KJ; (2017), 34(4) Carpenter J, Lee KJ; (2021), 38(4)
TG2: Selection of Variables and Functional Forms in Multivariable Analyses	https://www.stratos-tg2.org/home Perperoglou A, Heinze G, Sauerbrei W; (2018), 35(3) Heinze G, Perperoglou A, Sauerbrei W; (2021), 38(2)
TG3: Initial Data Analysis	https://www.stratosida.org/ Schmidt CO, Vach W, le Cessie S, Huebner M; (2018), 35(2) le Cessie S, Schmidt CO, Lusa L, Baillie M, Huebner M; (2021), 38(3)
TG4: Measurement Error and Misclassification	http://www.stratostg4.statistik.uni-muenchen.de/Home.html Freedman L, Kipnis V; (2018), 35(1) Shaw Pamela A, Boshuizen H; (2021), 38(1)
TG5: Study Design	https://tg5.stratosinit.org/home Gail MH, Cadarette S; (2019), 36(2)
TG6: Evaluating Diagnostic Tests and Prediction Models	Van Calster B, Steyerberg EW; (2020), 37(1) Steyerberg EW, van Calster B; (2022), 39 (2)
TG7: Causal Inference	https://www.ofcaus.org/home Waernbaum I, De Stavola B, Moodie E, le Cessie S, Goetghebeur E; (2018), 35(4) Waernbaum I, De Stavola B, Didelez V, Moodie E, le Cessie S, Goetghebeur E; (2022), 39(1)
TG8: Survival Analysis	Andersen PK, Abrahamowicz M, Therneau TM; (2019), 36(3)
TG9: High Dimensional Data	McShane L, Rahnenführer J; (2019), 36(1)
Simulation Panel	Boulesteix AL, Morris T, Sauerbrei W, Abrahamowicz M; (2020), 37(2)
Visualisation Panel	https://graphicsprinciples.github.io/
	Baillie M, Vandemeulebroecke M; (2020), 37(3)
Glossary Panel	Boeker M, Tippmann P, Day S, Huebner M, Sauerbrei W; (2020), 37(4)

In June 2019, the STRATOS initiative had its 2nd general meeting, held at the Banff International Research Station (BIRS, Canada). A key focus of this meeting was to initiate discussions about collaborative projects involving different TGs, as in practice observational studies pose many methodological challenges simultaneously that require expertise and guidance on different areas. We started discussions about collaborative projects involving different TGs. Details can be found in the [BIRS 2019 report on our website](#). Some results were presented in an invited STRATOS session at [IBC2022](#).

In parallel, STRATOS has started cross-TG discussions on some hotly debated issues that are important for the general statistical research community and for a large proportion of end-users of statistical methods. This includes the topic of ‘statistical versus machine learning techniques’ (Project leads: Rahnenführer and Lusa) and ‘use of statistical significance and p-values in real-world analyses’ (Project leads: Abrahamowicz, Carpenter and Kipnis). Resulting papers will attempt to honestly reflect internal discussions and potential diverging opinions on these highly controversial topics.

Panels to coordinate the initiative and to work on concepts to improve research

In the first BB article, in 2017, we listed ten cross-cutting panels, which were created to coordinate the activities of different TGs, deal with organizational issues relevant for all TGs, share best research practices, and disseminate research tools and results across the TGs. In 2018 the Visualization Panel was added, as visualization and a creative use of graphics can help at every stage of analysis, starting from the planning and the design of an experiment, and the very first data explorations, through to the communication of conclusions and recommendations. The STRATOS initiative embraces Open Science (OS) principles and endeavors to provide open access publications, illustrations with examples, and, if possible, publicly accessible datasets with programming codes, to promote transparency, utility, and reproducibility. Currently we are discussing how to facilitate the integration of OS principles in our panels and the research outputs.

In 2020, three panels (simulation, visualization and glossary) contributed short BB articles about how their activities help improve methodological research, their current work, as well as their main aims and plans for the future.

The Publication Panel and the Visualization Panel have the role of undertaking internal reviews of STRATOS manuscripts before submission, to help ensure that the initiative’s outputs follow best practice. Based on that experience, we propose that journals consider adding an expert for visualization to their editorial team. In the article from the Glossary Panel, we stressed the importance of establishing a widely accepted glossary of terms relevant for statistical analyses, which development will require support from relevant funded projects and/or seriously interested experts.

Simulation studies are the key instrument to compare statistical methods

In a tutorial paper entitled ‘Using simulation studies to evaluate statistical methods’, Morris (who joined STRATOS in 2018) et al. (2019) point out that ‘... simulation studies are often poorly designed, analyzed, and reported’ and then propose ‘... a structured approach for planning and reporting simulation studies, which involves defining aims, data-generating mechanisms, estimands, methods, and performance measures (“ADEMP”):

Another challenge of designing simulation studies is to ensure a fair comparison of alternative statistical methods. To achieve this, Boulesteix et al (2017a) proposed the concept of neutral comparison studies, which do not aim to demonstrate the superiority of a particular method (e.g., the one developed by the authors) and are thus not designed in a biased way that may (intentionally or unintentionally) create data structures that favor a specific analytical method. Led by Boulesteix, the STRATOS Simulation Panel (SP) published a related letter (Boulesteix et al (2017b)). A further logical step is to design and report simulation scenarios based on a variety of different, yet plausible, assumptions to assess if and how the relative advantages found for analytical methods depend on the assumed features of the data. In the BB article from the Simulation Panel, we provided further details and outlined ongoing and future projects to extend these two important concepts.

Guidance and education

In medical research, many data analyses are conducted by analysts with varying levels of statistical education, experience and interest. Consequently, it is not sufficient to aim at only providing guidance for experienced statisticians and experts. It may be even more important to provide some guidance for interested researchers with only basic training in applied statistics. The main aims of such guidance will be to (i) identify, and explain in a non-technical reader-friendly way, acceptable methods that are easily implemented and can provide valid responses to common analytical challenges, and (ii) highlight important weaknesses of some popular simple methods and clarify misconceptions regarding either their use for more complex data structures or misinterpretation of their results. Accordingly, to improve education, TGs and panels have various ongoing projects (e.g., a review of papers about statistical issues in medical journals, short videos, and shiny apps), but much more work is required. Some information can be found on the STRATOS website, as well as on TG websites.

External Collaborations

It is important for the STRATOS initiative to cooperate with other societies, initiatives and projects. In 2021 we signed a memorandum of understanding for collaboration between ISCB and STRATOS, and we will commemorate our 10th anniversary with a full-day mini-symposium at ISCB 2023. We are honored that IBS offered us the opportunity to write a series of short reports about our progress in the Biometric Bulletin.

Researchers from the European Organization for Research and Treatment of Cancer (EORTC) invited us to collaborate in the SISAQOL-IMI (Setting International Standards of Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials – Innovative Medicines Initiative; <https://www.sisaqol-imi.org/>) international multidisciplinary consortium, co-led by EORTC and the pharmaceutical company Boehringer Ingelheim. The consortium has been convened to generate recommendations to standardize the use, analysis, and interpretation of patient reported outcome (PRO) data in cancer clinical trials. Together with the pharmaceutical company Pfizer we lead work-package 3 ‘Feasibility of developing recommendations for non-RCTs, with single-arm studies as a case study’, which is led by Saskia le Cessie and Els Goetghebeur for STRATOS.

Currently, STRATOS has 100+ members from 20 countries on five continents. There are many challenging, important and interesting tasks to be tackled, and we welcome experienced new members (with sufficient expertise to contribute to a topic group or a panel)

as well as early career adjunct members who wish to participate in these activities (<http://www.stratos-initiative.org>).



Caption: Participants of the 2nd general meeting of the STRATOS initiative “Toward a Comprehensive Integrated Framework for Advanced Statistical Analyses of Observational Studies, June 2019, at the Banff International Research Station, Canada”. Authors of this article (from left): 1st row: Huebner (2), Sauerbrei (5), Shaw (6), Abrahamowicz (11); 2nd row: Gail (2), Keogh (11), 3rd row: Baillie (5). De Stavola could not attend. See the BIRS 2019 report for all participants and much more.

References

- Altman, D. G. (1994). The scandal of poor medical research. *BMJ*, 308 (6924), 283-284.
- Boulesteix, A.-L., Wilson, R., & Hapfelmeier, A. (2017a). Towards evidence-based computational statistics: Lessons from clinical research on the role and design of real-data benchmark studies. *BMC Medical Research Methodology*, 17, 138.
- Boulesteix, A. L., Binder, H., Abrahamowicz, M., & Sauerbrei, W. for the Simulation Panel of the STRATOS initiative (2017b). On the necessity and design of studies comparing statistical methods. *Biometrical Journal*, 60(1), 216-218.
- Kleinert, S. and Horton, R., 2014. How should medical science change? *The Lancet*, 9913(383), 197-198.
- Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38: 2074-2102
- Sauerbrei W, Abrahamowicz M, Altman DG, le Cessie S and Carpenter J on behalf of the STRATOS initiative. (2014) Strengthening Analytical Thinking for Observational Studies: the STRATOS initiative. *Statistics in Medicine*, 33: 5413-5432
- Sauerbrei, W., Bland, M., Evans, S. J., Riley, R. D., Royston, P., Schumacher, M., & Collins, G. S. (2021). Doug Altman: Driving critical appraisal and improvements in the quality of methodological and medical research. *Biometrical Journal*, 63(2), 226-246.
- Van Calster, B., Wynants, L., Riley, R. D., van Smeden, M., & Collins, G. S. (2021). Methodology over metrics: current scientific standards are a disservice to patients and society. *Journal of Clinical Epidemiology*, 138, 219-226.