Nicolas Savy*, Erica EM Moodie, Isabelle Drouet, Antoine Chambaz, Bruno Falissard, Michael R. Kosorok, Elizabeth F. Krakow, Deborah G. Mayo, Stephen Senn and Mark Van der Laan

Statistics, philosophy, and health: the SMAC 2021 webconference

https://doi.org/10.1515/ijb-2022-0017

Received February 1, 2022; accepted November 8, 2022; published online December 7, 2022

Abstract: SMAC 2021 was a webconference organized in June 2021. The aim of this conference was to bring together data scientists, (bio)statisticians, philosophers, and any person interested in the questions of causality and Bayesian statistics, ranging from technical to philosophical aspects. This webconference consisted of keynote speakers and contributed speakers, and closed with a round-table organized in an unusual fashion. Indeed, organisers asked world renowned scientists to prepare two videos: a short video presenting a question of interest to them and a longer one presenting their point of view on the question. The first video served as a "teaser" for the conference and the second were presented during the conference as an introduction to the round-table. These videos and this round-table generated original scientific insights and discussion worthy of being shared with the community which we do by means of this paper.

Keywords: artificial intelligence; Bayesian statistics; biostatistics; causality; health; philosophy.

1 Introduction

The Statistiques et Mathématiques Appliquées à la Cancérologie (SMAC), or Statistics and Mathematics applied to Oncology, group organizes annual SMAC conferences as part of the scientific programming of the Cancéropôle Grand Sud-Ouest, which is one of the seven Cancéropôle sections of the French National Cancer Institute (INCa). In an inter-regional and multidisciplinary approach, the GSO brings together nearly 500 teams of researchers and clinicians, mobilized against cancer over an entire territory, the Grand Sud-Ouest, covering the New Aquitaine and Occitanie regions. The Cancéropôle pools expertise and know-how to encourage and strengthen research projects, and accelerate therapeutic innovation and transfer of knowledge to the benefit

***Corresponding author: Nicolas Savy**, Toulouse Institute of Mathematics, University of Toulouse III and IFERISS FED 4142, University of Toulouse, Toulouse, France, E-mail: nicolas.savy@math.univ-toulouse.fr. https://orcid.org/0000-0001-9839-9598
Erica EM Moodie, Department of Epidemiology & Biostatistics, McGill University, Montréal, Québec, Canada, E-mail: erica.moodie@mcgill.ca
Isabelle Drouet, SND (UMR CNRS 8011), Sorbonne Université, Paris, France, E-mail: isabelle.drouet@sorbonne-universite.fr
Antoine Chambaz, MAP5 (UMR CNRS 8145), Université de Paris, Paris, France, E-mail: antoine.chambaz@u-paris.fr
Bruno Falissard, CESP, INSERM U1018, Université Paris-Saclay, Villejuif, France, E-mail: bruno.falissard@gmail.com
Michael R. Kosorok, Department of Biostatistics and Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, E-mail: kosorok@unc.edu
Elizabeth F. Krakow, Fred Hutchinson Cancer Research Center and University of Washington, Seattle, WA, USA, E-mail: efkrakow@fredhutch.org
Deborah G. Mayo, Department of Philosophy, Virginia Tech, Blacksburg, VA, USA, E-mail: mayod@vt.edu
Stephen Senn, Consultant Statistician, Edinburgh, UK, E-mail: stephen@senns.uk
Mark Van der Laan, Division of Biostatistics, School of Public Health, University of California, Berkeley, USA, E-mail: laan@berkeley.edu

of patients. Cancéropôle is structured along five scientific axes and several thematic clubs or working groups, one of which is the SMAC club. Further information can be found on the website www.canceropole-gso.org.

The SMAC club is open to all researchers working on population data, whatever their fields of application (epidemiology, psychology, social sciences, clinical research, etc.) as long as this work has applications in oncology. The objective of the SMAC club is to strengthen the links between these researchers, to help develop the skills of statistical teams and to make these skills better known to "users of statistics".

To achieve this goal, the club has, since its creation in 2012 in Toulouse, organized an annual conference known as SMAC days. The theme of the year is proposed by the local organizers at the Cancéropôle and the program is put together by the club's scientific committee. Topics explored during SMAC days are collected in Table 1. Further information on the SMAC club including the programs of previous SMAC days can be found on the club website.

The SMAC 2021 days were initially planned to be held in Toulouse in January 2021, and were motivated by a visit to the Mathematics Institute of Toulouse of Erica Moodie and David Stephens. Nicolas Savy (Mathematics Institute of Toulouse) and Erica Moodie (McGill University) co-organized the conference. The topics of the conference were in accordance with the expertise of the Canadian guests: causal inference and Bayesian statistics applied to medical research. In order to propose an original approach, an intersecting view between statistical and philosophical aspects of these questions of paramount importance was envisaged. The aim of this conference was thus to bring together not only statisticians and philosophers, but also doctors, epidemiologists, specialists in the humanities, . . . and to show that these disciplines can mutually enrich one another. The program was put together with the help of Olivier Claverie (Cancéropôle GSO) and Maël Lemoine (philosopher at Bordeaux University). The organizing committee was further enhanced by Isabelle Drouet (philosopher at Sorbonne University) whose expertise is precisely in philosophical aspects of Bayesianism and Causality.

Unfortunately, the COVID-19 pandemic forced organizers to move to a web conference format. Thanks to the Cancéropôle GSO team and especially Olivier Claverie, this event was a real success and its goal was achieved. Details of the program of the SMAC 2021 days can be found on the SMAC 2021 days website.

In Toulouse, SMAC days usually close with a round table engaging specialists with various points of view on a pre-specified question. In spite of the move to an online version of the workshop, the organizers wanted to maintain the round table in the program as a means of engaging speakers in a lively and interactive experience. To find a single relevant question was not easy given the breadth of topics covered at the conference. Thus, rather than having a single question to discuss, for the 2021 conference, numerous well-known scientists from different fields, with a recognized appetite for inquiry and a reputation for openness and depth of ideas, were solicited to formulate one or more key questions that may spark thought and debate. These scientists posed these key questions by means of 1–2 min "teaser" videos made available prior to the workshop, and the organizers asked participants to ponder these in the context of the themes addressed in the

Table 1: History of SMAC days conferences.

Year	Location	Title
2021	Online	Statistics, philosophy and health
2020	Online	Early phases, toxicities, competitive risks
2019	Bordeaux	Recent advances in joint models and the new challenge of immunotherapy clinical studies
2018	Toulouse	Personalized medicine, precision medicine
2017	Montpellier	Update on judgment criteria in oncology
2016	Bordeaux	Biostatistical and biomathematic modeling of imaging data
2015	Toulouse	Modeling and simulation of clinical trials
2014	Montpellier	Assessment and analysis of the quality of life in oncology, new methodological developments
2013	Bordeaux	Dynamic predictions for repeated markers and repeated events
2012	Toulouse	Mixed hidden markov models and treatment of cancer cohort data

workshop – mainly Bayesian inference, causality, and Artificial Intelligence. The scientists further developed their thoughts on the key questions and sometimes proposed solutions in a 6–7 min video that was broadcast during the conference. These videos provided a rich opening to the round table that followed.

Several contributions will be published in this special issue of *The International Journal of Biostatistics* together with this article, which chronicles the concluding round table. It was important for the organizers to document this concluding event and to share not only the format of this novel discussion but also the exciting ideas that were put forward. The organizers would like to warmly thank *The International Journal of Biostatistics* for agreeing to publish such a special issue.

This paper is divided into two main sections. The first section is a compilation of the contributions of scientists who accepted to play the game of the "video-captured ticket of the mood" on the topic of Bayesian statistics, causal inference, and philosophy in health. For each contribution, a short biography of the author is given together with a summary of the ideas developed in the video written by the contributor and the link to the videos. This section is thus divided into subsections, each titled by a key question posed by the contributors. The organizers then provide a report on the main ideas developed during the discussion in the concluding round table; this forms the next section of this paper. The video of the round table is not available, due to difficulties in obtaining the authorization of all of the participants. Finally, we close with some brief concluding remarks.

2 Summaries of invited contributors

2.1 Do you think that statisticians should learn to rely on deterministic models?

We open with the contribution of Professor Antoine Chambaz, professor of statistics at Université de Paris, a member of MAP5 (UMR CNRS 8145) and the head of its Statistics team. He is also the director of FP2M (FR CNRS 2036), the Parisian Federation of Mathematical Modelling. Professor Chambaz holds a PhD in mathematics (2003) from the Paris-Sud University. His main research interest is in theoretical, computational and applied statistics, and in causality. In particular, he contributes to the development and application of the targeted learning methodology, with a focus on applications to medicine and precision medicine, and studies problems at the intersection of statistics and machine learning. Professor Chambaz coedits *The International Journal of Biostatistics*. Professor Chambaz's video can be seen at this link.

To answer questions on a phenomenon of interest, the statistician defines a statistical model $\mathcal{M}$ for the law P of the experiment consisting of observing the phenomenon, and learns some problem-specific, relevant features of P. In parallel, countless deterministic models have been developed to model natural phenomena. What do I call a "deterministic model"? It is a mathematical model obtained by combining elementary blocks, each block built based on the fundamental laws of physics (e.g., conservation of mass-energy, linear momentum). The model provides a (possibly multi-scale) description of a phenomenon under the form of equations involving quantities that are observed or not. Physico-chemical, physiological, environmental parameters drive the regime of the model. Such deterministic models lend themselves to being combined in sophisticated "hierarchies of models". Here are my questions to my esteemed colleagues: do you also think, as I do, that statisticians should learn to rely on deterministic models? If so, how?

2.2 What do statistical tests really say?

What follows is the contribution of Professor Bruno Falissard. After some initial training in mathematics and fundamental physics, Professor Falissard engaged in medical studies and specialized in Child and Adolescent Psychiatry in 1991. He completed a PhD in biostatistics and was an assistant professor in child and adolescent psychiatry in 1996–1997, an associate professor in biostatistics at University Paris-Nord in 1997–2002, and a full professor in biostatistics at University Paris-Sud from 2002. Professor Falissard is the head of the Center

of Epidemiology and Population Health and co-author of more than 400 papers. His areas of research are in methodology and epistemology of medical research, in particular in psychiatry. Professor Falissard's video can be seen at this link.

Between 1920 and 1930, the decade that can be considered as the origin of the formalization of statistical tests, a drama took place. Among the characters, there was a "father" (Karl Pearson), a "son" (Egon Pearson), a "genius" (Ronald Fisher) and a "mathematician" (Jerzy Neyman). This drama could appear at first glance as a ridiculous family dispute: is "p" or "alpha" the cornerstone of statistical inference? It is in fact much more than that, and a hundred years after the statistical community is still burdened with very simple questions: is acceptance of a null hypothesis possible? Are one-sided tests conceivable? Or, on the contrary, are statistical tests always one-sided? Is there really a difference between "$p = 0.048$" and "$p = 0.052$"? What did R. Fisher and J. Neyman mean, exactly, when claiming that the former dealt with inductive inference and the latter with inductive behaviour?

2.3 Is the field of data science ready to solve the world's problems?

We now turn to the contribution of Professor Michael R. Kosorok. Professor Kosorok is a W.R. Kenan, Jr. Distinguished Professor of Biostatistics and Professor of Statistics and Operations Research. He is the 2019 ASA Gottfried E. Noether Senior Scholar Award recipient, and is a Fellow of the American Association for the Advancement of Science, the American Statistical Association, and the Institute of Mathematical Statistics. His video can be seen at this link.

The field of data science, an amalgamation of many disciplines, including the foundational fields of computational sciences, informatics, mathematics, and statistics, is crucially important in the quest to improve the world through wise use of data and information. Each of the foundational constituent disciplines brings crucial perspectives and capabilities to the table, but are we flexible and deep enough in the way we are thinking about the foundations of data science? Are we working sufficiently well with applied domain experts and key stakeholders in society to maximize benefit and minimize harm? Are we inclusive enough in how we collaborate with application domain experts?

The focus of the field of statistics is inference, which, generally speaking, refers to the question of how conclusions from data will have meaning in the real world. Statistics has a long tradition of models and frameworks for answering these questions; however, we are at the cross-roads of a potential major transition in the field. The essence of this question at the cross-roads is whether we are willing to sufficiently adapt as a field to rise to the occasion that presents itself. These questions include the following four: (1) can we expand our theoretical foundations, methods and applications to be sufficiently broad to handle all data and analytic situations which may present themselves? (2) Are we ready to better teach and train the world on the fundamentals of inference, inferential intuition, and practical quantitative reasoning, including teaching other data scientists, domain science experts, policy makers, and society at large? (3) Are we prepared to work with new sources of data and information, including qualitative, non-stochastically generated samples, and data acquired in real time, for example? (4) Are we willing to take on collaborative leadership to a greater degree, as needed? Similar questions arise in the other constituent foundational data science domains, and we all need to proactively prepare for the coming explosion in data science and its applications.

2.4 Why can't we cure more patients who have advanced cancer?

We now have the contribution of Dr Elizabeth Krakow. Dr Krakow is an assistant professor in the Clinical Research Division at Fred Hutchinson Cancer Research Center and in Medical Oncology at the University of Washington. Her clinical practice is focused on blood and bone marrow transplantation. She conducts clinical trials of immunologic therapies to prevent and treat leukemia relapse after such transplantations. Her video, taking the format of a conversation with Dr Jerald Radich also of the Fred Hutchinson Cancer Research Center, can be seen at this link.

Biologists have developed "evolutionary" models of cancer progression based on relationships among competing and cooperating malignant clones, the patient's immune system, treatments administered, and tissue microenvironments. They have characterized basic mechanisms of tumour progression under the pressure exerted by anti-neoplastic treatments, such as: Darwinian selection pressure, culling-the-herd (where all malignant clones are reduced proportionately, only to regrow once treatment is stopped), and stochastic emergence of new mutations. However, there is a gap when applying these models in clinical practice. We are limited by the cost of single cell sequencing and by sampling bias introduced when a small volume of cancerous tissue is excised at a single time-point (because serial, large-volume tissue biopsies are usually too invasive, and the cancer continues to change after the biopsy is taken). Therefore, when looking at a given patient, we usually do not know which of these models applies to his or her cancer progression. Using less costly bulk tumour sequencing on sub-optimally sized tumour samples, how can we infer which phylogenetic branches of clonal evolution are being traced by a specific patient's cancer? How can we use these biological models to guide treatment selection?

A second question arises from the following challenge: clinical trials are generally designed by frequentists, but then we, as clinicians, try to be Bayesian. We set up some prior probability of survival for the patient based on our experience, tumour staging, etc. and then, based on the clinical trial results, we estimate how the probability of survival will change if we choose treatment A or treatment B. But how can we make the necessary leap from trial-derived statistical models to patient-specific prediction? Moreover, for many cancer-associated mutations or combinations of mutations, the population frequency is so low that it is impossible to launch a traditional randomized treatment trial. The tremendous inter-patient variability in drug metabolism and immunologic responses adds further complexity. Can society build the infrastructure to collect and fairly disseminate big, yet granular and accurate, data sets that might support new clinical study designs that take advantage of our molecular-level understanding of cancer evolution and that are better suited to developing and validating highly personalized adaptive treatment strategies?

2.5 Four questions to ask in scrutinizing accounts of statistical inference

Next, we have the contribution of Professor Deborah Mayo. Professor Mayo is Professor Emerita in the Department of Philosophy at Virginia Tech. She is author of Error and the Growth of Experimental Knowledge (1996, Chicago) which won the 1998 Lakatos Prize. Her most recent book is Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars ([1], CUP). She blogs at errorstatistics.com. Her video contribution can be seen at this link.

The main source of today's crisis in replication is that high-powered methods make it too easy to uncover impressive looking findings, even if they are false. If little or nothing has been done to uncover flaws in inferring a claim, we do not have evidence for it. It has not passed any kind of a severe test. So the first question to ask in scrutinizing a statistical account is: does it permit violations of this minimal requirement for evidence? A second question concerns presuppositions about the role of probability in inference. In frequentist error statistics, probability is intended to assess and control the probabilities of misleading interpretations of data. To those who assume probability should provide degrees of belief or probability in hypotheses, a p-value appears to be irrelevant, or necessarily misconstrued. But that is to presuppose a rival philosophy of inference akin to Carnap insisting Popper's falsificationist become a confirmationist. So the second question to ask is: does an account enable falsifying claims at least statistically? We often hear that it is just too easy to obtain small p-values. Yet replication attempts find it is difficult to get small p-values when they have preregistered hypotheses. What this shows is that the problem is not p-values, but failing to adjust their values for cherry picking, multiple testing, and other biasing selection effects. A third question is: is the uncertainty measure directly altered by biasing selection effects, as we think they should be? The same data-dredged hypothesis can occur in methods put forward as alternatives to p-values: likelihood ratios, Bayes' factors, and others, except now we lose control of the grounds to criticize inferences for flouting error control. Finally, even those who criticize p-values will employ them, at least if they care to test the assumptions of their statistical model – an altogether important task. So a fourth question is: is the account capable of testing its own assumptions?

So, if someone is selling you an account or a reform, where the answer to any of these questions is no, you may wish to hold off buying it.

2.6 Being wrong about randomisation

The next contribution was that of Professor Stephen Senn. Professor Senn is a medical statistician. He has worked in England, Scotland, Switzerland, Luxembourg, and France as an academic and as a research scientist, and also in the pharmaceutical industry. His research interests are statistical methods in drug development. Professor Senn's video contribution can be seen at this link.

You sometimes hear the following criticism of randomisation: whether or not some factors are balanced as a result of random allocation, since there are indefinitely many factors, some of them must be imbalanced. Therefore, the analysis of randomised clinical trials is not valid. There may be good arguments against randomisation but this is a really bad one. For me, it is a litmus test statement. If you believe it, I know that you do not understand randomization. I also strongly suspect that you do not understand the basics of statistical analysis. And I am absolutely confident that you have never tried to simulate the problem that you claim exists. Doing is a cure for stupid speculation.

To show how wrong it is, let me give you a statement that is true and that is almost the opposite of this. If we knew that every possible prognostic factor were balanced, then the standard analysis of clinical trials would be wrong. The critics of randomisation have picked up the problem from the wrong end. They have started by assuming that the goal is perfect estimation. But perfect estimation cannot accept any imperfections. Therefore, we must always worry about things that might be imbalanced. It is clear that randomisation cannot balance everything. Therefore, randomisation cannot be trusted. If you believe that this is the way to start the argument, let me ask you a question. Do you think that anything useful can be said about the probability that a fair die will show three sixes if rolled three times? After all, there are indefinitely many factors that could affect the outcome.

The correct way to start is to accept that perfect estimation is impossible. Setting probabilistic bounds on how imperfect your estimates are is possible and useful. It is this goal that randomisation supports. Standard statistical methods make an allowance for random imbalance. If we knew that every possible prognostic factor were balanced, then the standard analysis of clinical trials would be wrong.

2.7 The need for realistic statistical models and utilization of sieve MLE (highly adaptive lasso)

Finally, the videos concluded with the contribution of Professor Mark van der Laan. Professor van der Laan is the Jiann-Ping Hsu/Karl E. Peace Professor of Biostatistics and Statistics at the University of California, Berkeley. He has made contributions to survival analysis, semiparametric statistics, multiple testing, censored data and causal inference. He also developed the targeted maximum likelihood methodology and general theory for super-learning. He is a founding editor of the *Journal of Causal Inference* and *The International Journal of Biostatistics*. He has authored four books on targeted learning, censored data, and multiple testing, authored over 350 publications, and graduated 55 PhD students. He is the recipient of the Mortimer Spiegelman Award (2004), the COPSS Presidents' Award (2005), and the van Dantzig Award (2005). Professor van der Laan's video can be viewed at this link.

An important question raised is whether our field is willing to adopt the formulation of statistical estimation problems in terms of (1) realistic statistical models for the data distribution, and (2) estimands as defined as mappings from these possible data distributions into the answer to question of interest (or approximation thereof). The emphasis is on the statistical model being realistic so that it only incorporates assumptions about the data generating experiment that are known to be true of at minimal can be well defended. As a result of accepting this translation of the real world into a statistical estimation problem, is one then willing to define an *a priori* defined estimator, possibly based on outcome blind data, and corresponding statistical inference, augmented possibly with a sensitivity analysis with respect to non-testable assumptions

that enhance the interpretation of the estimand of the data distribution. We often refer to this as the roadmap for statistical learning and causal inference [2, 3].

The second issue is related to this so-called formulation of a statistical estimation problem in the sense that it requires utilization of machine learning algorithms that are able to learn the true target functions of the data distribution at a fast enough rate. We conjecture that good robust machine learning algorithms should aim to be sieve maximum likelihood behaving estimators so that they solve a rich set of score equations, but also generate scores that are embedded in a well-behaving Donsker class. In particular, a more recent nonparametric maximum likelihood estimator terms the Highly Adaptive Lasso MLE (HAL-MLE) is contrasted to other machine learning algorithms [4, 5]. It is noted that HAL-MLE indeed solves a class of score equations that approximates the space of all score equations as sample size increases, and, in addition, it converges at a $n^{-1/3}$ rate to the true target function, up till log n-factors. As a result it has been shown to result in efficient plug-in estimators of a large class of target features of the data distribution [6]. The key properties which makes this HAL-MLE so robust are, in particular, that it solves a rich set of score equations and that the linear span of these score equations remains embedded in a nice Donsker class with excellent entropy integral. We suggest that this perspective on benchmarking machine learning algorithms would move the field forward in the right direction by focusing on algorithms that satisfy these key fundamental properties that drive their statistical behaviour.

3 The closing rountable

Erica Moodie moderated the round table, noting the progression of ideas from Elizabeth Krakow setting the clinical scene for precision medicine in oncology, to questions of design raised by Stephen Senn, proposals in a framework of estimation put forward by Mark van der Laan, and the importance of estimating and replication highlighted by Bruno Falissard and Deborah Mayo. Antoine Chambaz suggested a template for analysis, while Michael Kosorok asked a question at the heart of statistics and philosophy in health: whether data science can solve the challenges we face as a society.

Sander Greendland continued on the discussion on replication, focusing on sociological incentives to find publishable results, noting that this could include not only 'p-hacking' to obtain statistically significant results, but can also cover 'CI-hacking' in which variance is inflated to avoid a statistically significant finding, which could be noteworthy if it contradicts current knowledge or recent significant findings. He also presented several psychological biases to which he attributed various misinterpretations of statistical models or results. Philosopher Jon Williamson raised the question of whether the so-called replication crisis was indeed so problematic, whether it was unsurprisingly that 'new' findings fail to replicate. He suggested that might be more concerning of results considered truly established failing to replicate. Deborah Mayo raised the point that although Neyman and Pearson had emphasized decision-making, there was never an expectation that these methods would be used bluntly but rather in a more subtle way such that p-values just either side of the 0.05 cut-off would not have been treated differently. Professor Mayo went on to praise the importance of redefining significance and the use of the Bayes' factor to bring more conscious understanding to the decision-making side of inference.

Dr Krakow was asked how the question of replication arises in the context of oncology, and specifically in the setting of bone marrow transplants where the number of patients (and thus potential trial participants or study subjects) is limited. Dr Krakow returned to the 'culture' that Professor Greenland had mentioned, where team science is often not rewarded to the same extent as sole researcher endeavours, even though the latter tend to produce smaller, less conclusive findings. She advocated for teaching the sociology of statistics and most importantly some psychology of reasoning – cognitive biases that may lead to analytic choices that bias findings – at an early stage in scientific and clinical training. Professors Mayo and Greenland then discussed the merits of thresholding in decision-making, and what solutions might exist to the problems it raises. Professor Greenland raised, as a possibility, the reviewing of articles with results withheld so that the

methods and data can be evaluated in the absence of knowledge of whether a given p-value threshold was crossed.

Nicolas Savy then posed the question to David Stephens of whether the future of statistics will be most concerned with the massive size of data or the evolution of analysis techniques, noting that these two issues may not be in conflict. Professor Stephens responded that there is always a risk that the methods development in our 'mathematics laboratories' are too simplistic for the data being created and collected in the real world. He noted that this leads him to stress in his work the importance of representing and reporting uncertainty, although he recognized that doing so may not be straightforward. Professor Stephens highlighted the idea from Mark van der Laan's video, noting that the use of principled and flexible models offers a greater possibility of reflecting reality, even if some aspects are not verifiable.

Dr Krakow then asked Antoine Chambaz to elaborate on how deterministic models might bring together biological knowledge and clinical practice with statistical modelling. Professor Chambaz said that progress has been slow, but that there is much promise in the deterministic models that can precisely describe biological mechanisms that may give greater predictions. Dr Krakow replied that by year's end, more than 15% of Americans' health data will be in the hands of private industry, with the intention of mining these data to provide decentralized decision-making. She noted, however, that treatment-recommendation algorithms that cannot be explained or understood raise ethical and legal questions for clinicians who may choose to rely on them. Dr Krakow noted that the scale at which treatment recommendation algorithms driven by artificial intelligence operate is far greater than that of a single physician working with their patient pool, and thus the potential for harm is far greater if the algorithms perform badly; this was her motivation for wondering to what extent prior biological knowledge and deterministic models could be incorporated into such learning algorithms to ensure greater safety. Professor Moodie raised the example of existing warfarin dosing, where nearly 20% of the time, physicians opted not to use the recommended dose. She noted that physicians tend to know their own patients very well, but also know only a relatively small number of patients – thus physicians often base decisions on highly detailed information from modest numbers of individual patient profiles, whereas the algorithms are often relatively simplistic (possibly unrealistic) models based on a far greater sample size. Professor Chambaz agreed that this goal of precision medicine is being tackled by many: the goal is to learn from averages and deviations from the average to subsets or even single individuals. He noted that we have fantastic theoretical results, but large gaps remain between the theoretical and the real-world use of such approaches. Dr Krakow reiterated earlier sentiments of Professor Stephens', that there is a 'leap of faith' required to move from statistical models to clinical practice.

The round table discussion was energetic and engaged, with active participation from all of the pillars of the 2021 SMAC days – the statisticians, the philosophers, and the oncologists alike. The overlap in research goals and eagerness to collaborate and draw on mutual interests and strengths was evident, and a strong indication of the success of the workshop. The round table concluded with a word of thanks to all participants and the session organizers.

4 Concluding remarks

The questions posed and ensuing discussion highlighted important themes in our field. It is clear that the question of how to ensure reproducibility, through design and by statistical method of decision making, is still a point of some contention. There is fierce agreement on the importance of making meaningful inference and decisions, and the importance of publishing scientific findings regardless of the result (i.e., whether or not a threshold of statistical significance was achieved). However even if decision-making were a settled question, the issue remains of how faithful statistical models are to the real-world, and whether the data-driven complexity of machine learning models can assure either fidelity to the real world or conformity to the ethical and legal requirements of clinical care. These questions require the attention of statisticians, but it is clear that the answers are not to be found in statistics alone, but rather through close collaboration with

clinicians, philosophers and ethicists, and perhaps also legal experts. Finally, accurate and honest reflection of uncertainty should not be overlooked.

The success of the roundtable was a testament to that of the whole SMAC conference. Although it was held online, the workshop gave rise to inspiring discussions and committed debates. Pluri-disciplinarity undoubtedly contributed to the richness of the event and in turn all the participants benefited from the views of scholars outside their area of specialization. The organizers would like to underline the academic importance of such events. They warmly thank all the participants and especially the participants in the roundtable, who accepted to play the game of formulating questions and elaborating videos. They are proud to share this publication with them. The Cancéropôle GSO should also be warmly thanked for its financial, technical and scientific support, with special credit to Olivier Claverie's unfailing investment in this project.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: Nicolas Savy has benefited from the support of INCa-Cancéropôle GSO – AAP 2020 Emergence of projects, and from the Tremplin 2021 program of the University of Toulouse 3. Erica EM Moodie is a Canada Research Chair (Tier 1) in Statistical Methods for Precision Medicine and acknowledges the support of a chercheur de mérite career award from the Fonds de Recherche du Québec, Santé.
Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References and Suggested readings [2–13]

1. Mayo D. Statistical inference as severe testing: how to get beyond the statistics wars. Cambridge: Cambridge University Press; 2018.
2. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. Epidemiology 2014;25:418–26.
3. van der Laan MJ, Rose S. Targeted learning: causal inference for observational and experimental data. Springer series in statistics. New York: Springer; 2011.
4. Benkeser D, van der Laan MJ. The highly adaptive Lasso estimator. Proc Int Conf Data Sci Adv Anal 2016;2016:689–96.
5. van der Laan MJ. A generally efficient targeted minimum loss based estimator based on the highly adaptive lasso. Int J Biostat 2017;13:10.
6. van der Laan MJ, Benkeser D, Cai W. Causal inference based on undersmoothing highly adaptive lasso 2019. In: Proceedings of the symposium on beyond curve fitting: causation, counterfactuals, and imagination-based AIAAAI spring 2019. Stanford, CA; 2019.
7. Fisher RA. Statistical methods and scientific induction. J Roy Stat Soc B 1955;17:69–78.
8. Fisher RA. Statistical methods for research workers, 14th ed.—revised and enlarged. New York: Hafner Publishing Co.; 1973.
9. Marks HM. Rigorous uncertainty: why RA Fisher is important. Int J Epidemiol 2003;32:932–7.
10. Mayo D. P-values on trial: selective reporting of (best practice guides against) selective reporting. Harvard Data Sci Rev 2020;2. https://doi.org/10.1162/99608f92.e2473f6a. https://hdsr.mitpress.mit.edu/pub/bd5k4gzf/release/4. In this issue.
11. Merlo LM, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. Nat Rev Cancer 2006;6:924–35.
12. Senn S. Seven myths of randomisation in clinical trials. Stat Med 2013;32:1439–50.
13. Senn S. Randomisation is not about balance, nor about homogeneity but about randomness. Error Stat Philos 2020. https://errorstatistics.com/2020/04/20/s-senn-randomisation-is-not-about-balance-nor-about-homogeneity-but-about-randomness-guest-post/. In this issue.