

# BIOMETRIC BULLETIN

International Biometric Society Internationale Biometrische Gesellschaft Société Internationale de Biométrie

*"Biometry, the active pursuit of biological knowledge by quantitative methods."* - R.A. Fisher, 1948

## President's Corner



I hope you are all safe. At the time I was writing my previous contribution, the COVID-19 pandemic had just started and was primarily contained within Asia. The Officers were still confident that

the crisis would be over by July and that we would be able to hold our IBC 2020 meeting as scheduled in Seoul, South-Korea. In the meantime, we all know that the virus is spreading over the entire world and all continents have been affected already. Many countries had to go into lockdown, closed their borders, and imposed major restrictions on people moving around and meeting each other. Suddenly, teleworking was the norm, shops, bars and restaurants closed, and hospitals had to prepare to treat many severely ill affected by the virus. It is clear that the pandemic has resulted and still is resulting in many victims worldwide. So, my first concern is that you, your relatives, and your friends are safe and well.

COVID-19 not only has a major medical and societal impact, it also affects our professional life. Meetings need to be held online. On campus teaching is no longer possible and we have to develop distance learning programs. Different formats for student evaluation need to be explored. Many of us are now heavily involved in modeling COVID-19 data such as the number of new infections, the number of hospitalizations, the number of people in intensive care units, and the number of deaths. Others help clinicians set up

clinical trials conducted in an effort to evaluate potential treatments and vaccines. The main challenge is that there is a lot of time pressure, there are many potential treatments, and the number of new cases starts to rapidly decrease due to the severe lockdown measures governments had to impose. Also, many statisticians, data analysts, and epidemiologists are called upon to inform the general public about how to interpret data collected, to explain the potential impact of certain measures, or to advise local governments on how to relax the lockdown measures. This way we can show clinicians and policy makers how sound statistical methodology can help to correctly interpret epidemiological data, to predict the impact of exit strategies, and to design studies which use the available patients as efficiently as possible to evaluate the available potential therapies.

COVID-19 also has a major impact on IBS. On March 18 we had a virtual Executive Board meeting. While that was still at the beginning of the spreading of the virus in many parts of the world, it was already clear that IBC could not be held as planned in July 2020. The Board extensively discussed alternatives and the officers were asked to explore the possibility to postpone the meeting to August 2021, when the venue would still be available. While postponing initially seemed the most obvious choice, we soon realized this option was not feasible. First and foremost, the course of COVID-19 is unclear and many countries already announced that travel restrictions will be retained for a long time. Even if many countries were to lift their travel restrictions in time, it is unclear how many of our typical participants would be

*Continued on p. 6*

## IN THIS ISSUE

President's Corner .....	1
Region Key .....	2
From the Editor .....	3
Response to the Editor - Basic Theme IV: Inductive and Deductive Reasoning .....	3
Basic Themes V & VI (Reader's responses are invited) .....	4
International Biometric Conference   Virtual Learning Series .....	7
IBS Awards Update .....	7
Editorial Update .....	8
Software Corner .....	10
STRengthening Analytical Thinking for Observational Studies (STRATOS) .....	11
Region News .....	13
* Eastern North American Region (ENAR) .....	13
* French Region (SFB) .....	13
* Japanese Region (JR) .....	14
* The Netherlands Region (ANed) .....	14
* Spanish Region (REsp) .....	15
* Western North American Region (WNAR) .....	15
Announcements & Upcoming Events ...	15
Get Connected .....	16

## BIOMETRIC BULLETIN

ISSN 8750-0434

Copyright © 2020

International Biometric Society

Biometric Bulletin will be published four times a year in March, June, September, December.

International Biometric Society

1120 20th Street, NW, Suite 750 • Washington, DC 20036-3441 USA

Telephone: +1 (202) 712-9049 • Fax: +1 (202) 216-9646

Email: [ibs@biometricsociety.org](mailto:ibs@biometricsociety.org) Website: <http://www.biometricsociety.org>

The Biometric Bulletin is available to members of the Society as part of their annual dues. The views of contributions to this publication should not be ascribed to the International Biometric Society. Reproduction for commercial purposes is allowed if the source is acknowledged

### Editor

Dr. Ajit Sahai, Ex. Founder-Professor & Chair,  
Department of Medical Biometrics; JIPMER, Pondicherry; India  
EMAIL: [editor.bb2019@gmail.com](mailto:editor.bb2019@gmail.com)

### Associate Editor

Garth Tarr, University of Sydney, Australia; EMAIL: [garth.tarr@sydney.edu.au](mailto:garth.tarr@sydney.edu.au)

### Regional Correspondents

Joanna in 't Hout (ANed), Rina Parrry (AR), Elizabeth Williamson (BIR) Hein Putter (Channel Network), Reinhard Vonthein (DR), Dr. Daniel Nevo (EMR), Jarcy Zee (ENAR), Olayemi Oluwasoga (GNi), Zofia Hanusz (GPol), Iain McDonald, (GSAf), Wellington Mushayi (GZim), Ikuko Funatogawa (JR), Esa Läärä (NR), Sophie Vanbelle (RBe), Wagner Bonat (RBras), Novie Younger (RCAC), Anabel Forte (REsp), Robert Faivre (RF), Livio Finos (RItl), Sangbum Choi (Rko), Dominik Heinzmann (ROeS), Henry G. Mwambi (SUSAN), Megan Othus (WNAR), Jialiang Li (SING), Dan Kajungu (GUgan), Agnes Ankomah (GGha), John Mwangi (GKe), Njoku Ama (GBot), Cristian Meza (GCl), Katja Ickstadt (CEN), Alia Sajjad (PKSTAN), Yan You (CHINA), Gema Alexandra Zambrano (ECU), Jupiter Simbeye (GMal)

### International Biometric Society Executive Board

President: Geert Verbeke, Belgian Region

Outgoing President: Louise Ryan, Australasian Region

Secretary/Treasurer: Vicente Núñez-Antón, Spanish Region

Directors: Karen Bandeen-Roche, United States; Frank Bretz, Switzerland; Mark Brewer, United Kingdom; Elizabeth Brown, United States; Brian Cullis, Australia; Krista Fischer, Estonia; Joel Greenhouse, United States; Freedom Gumedze, South Africa; Tae Rim Lee, Korea; José Pinheiro, United States; Luzia Trinca, Brazil; Andreas Ziegler, Germany

### Biometrics Editors

Geert Molenberghs (Executive Editor), Biostatistics & Statistical Bioinformatics Hasselt University & K.U. Leuven I - Biostat; EMAIL: [geert.molenberghs@uhasselt.be](mailto:geert.molenberghs@uhasselt.be)

Debashis Ghosh, Penn State University of Biology - Wartik Lab; EMAIL: [ghoshd@psu.edu](mailto:ghoshd@psu.edu)

Alan Welsh, ANU College of Business and Economics, Email: [Alan.Welsh@anu.edu.au](mailto:Alan.Welsh@anu.edu.au)

Mark Brewer, BioSS; EMAIL: [markb@bioSS.ac.uk](mailto:markb@bioSS.ac.uk)

Chantal Brodie and Ann Hanhart (Editorial Manager), Email: [biometrics@biometricsociety.org](mailto:biometrics@biometricsociety.org)

### Journal of Agricultural, Biological and Environmental Statistics (JABES) Editor

Brian Reich, Department of Statistics, NC State; EMAIL: [bjreich@ncsu.edu](mailto:bjreich@ncsu.edu)

### International Business Office

Executive Director: Peter Doherty, CAE

Director of Operations: Kristi Wolford

Membership and Travel Awards: Amanda Bignall

Membership Coordinator: Starr Ellis

Staff Support: Daniel Medina

Content Marketing Manager: Cat Wilson

## Region Key

### Regions

RArg - Argentinean Region  
AR - Australasian Region  
ROeS - Austro-Swiss Region  
RBe - Belgian Region  
GBot - Botzwanian Region  
RBras - Brazilian Region  
BIR - British and Irish Region  
RCAC - Central American-Caribbean Region  
GCl - Chilean Region  
CHINA - Chinese Region  
EMR - Eastern Mediterranean Region  
ENAR - Eastern North American Region  
ECU - Ecuadorian Region  
GEth - Ethiopian Region  
RF - French Region  
DR - German Region  
GGha - Ghanian Region  
IR - Indian Region  
RItl - Italian Region  
JR - Japanese Region  
GKe - Kenyan Region  
Rko - Korean Region  
GMal - Malawi Region  
GNi - Nigerian Region  
NR - Nordic-Baltic Region  
PKSTAN - Pakistani Region  
GPol - Polish Region  
GRo - Romanian Region  
SING - Singaporean Region  
GSaf - South African Region  
REsp - Spanish Region  
ANed - The Netherlands Region  
GUgan - Ugandan Region  
WNAR - Western North American Region  
GZim - Zimbabwean Region

### Networks

CEN - Central European Network  
CN - Channel Network  
EAR - East Asian Network  
SUSAN - Sub-Saharan Network

## From the Editor

It is of interest to note that during a series of lockdowns simultaneously imposed in multiple countries, an enhanced online exchange of thoughts within and between the coherent groups like ours proved to be a good alternative to keep us updated. Although, this time the reporting from regional correspondents may be affected, we have been successful in continuing with our IBS regional activities to a great extent despite several constraints. In the meantime the webinars and virtual meetings have come-up in a big way to rescue continuance of our regional as well international activities. The IBS Committee on communications had recently arranged a Zoom meeting this April to ensure our outreach mission's progress is uninterrupted. The regular updates received timely from Biometrics & JABES; as well the contributions from Software-Corner and STRATOS deserve much appreciation. Considering the need of the hour a detailed overview analysis undertaken by Prof. Geert Verbeke on how the current situation had already affected IBC2020 and further what alternatives are still feasible, explored or being planned are well shared by him in 'President's Corner'.

While a write-up received from Prof Kuldeep Kumar, Australia in response to the latest basic theme 'inductive and deductive reasoning' is being published under the column 'Response to the Editor', another critical urge clubbed with it was to invite a meaningful debate on 'Evidence Based Medicine' that somehow had missed the attention and should takeoff soon. Considering the fact that only a few regions have reported this time I am motivated to submit the write-ups on two themes instead of one, i.e. 'Basic-theme:V on 'Validity'; and theme VI on 'Error and Bias'; additionally their strong bondage with the concepts of 'Consistency, Reliability, Sensitivity, Precision, Accuracy, Specificity & Efficacy' are also very briefly outlined. It is well understood that these efforts are with a clear objective of inviting responses from the readers by provoking their original fundamental thoughts behind each of these loaded terms. Otherwise, there already exists a body of organized scientific knowledge on these concepts acquired over its due course and well written in the forms of chapters and books on the related subjects such as; psychology, education, philosophy, sociology and the basic research-methodology prevailing overall. Biometrics by default is focused on measurements in life sciences diagnostics, actions, effects and predictions and therefore the debate should continue around it. Of course, the next basic theme in continuity and logical sequence should be 'Evaluation of Efficiency, Efficacy and Effectiveness' of large scale projects & programs to be covered in the next issue.

**Ajit Sahai**

Biometric Bulletin, Editor

## Response to the Editor - Basic Theme IV: Inductive and Deductive Reasoning

(Kuldeep Kumar, Bond University, Australia)

The Editor has raised a very timely and interesting debate on inductive and deductive reasoning. However, I am slightly deviating from this issue, and I will rather concentrate on a seminal paper by Leo Breiman (2001) on "Statistical Modelling: The Two Cultures" and later on I will try to link it to this topic. According to Breiman, "There are two cultures in the use of statistical modelling to reach conclusions from data. One assumes that the data are generated by a given stochastic data model. The other uses algorithmic models and treats the data mechanism as unknown". In the first case like regression models, logistic regression, and Cox-model, the values of the parameters are estimated from the data and the models are used for information and/or prediction. The second case, which Breiman calls "The Algorithmic Modelling Culture", the analysis in this culture is complex and unknown. Their approach is to find a function  $f(x)$  through an algorithm that operates on  $X$  to predict  $Y$  and Breiman himself called this a black box. The machine learning algorithms like Decision Tree, Random Forest, and Stochastic Gradient Boosting and to some extent ANN falls in this category.

I am sure a lot of biometricians believe only in statistical modelling, but some of them also believe in machine learning and few of them try to bridge the gap between the two (machine learning and statistical modeling). I am a big fan of Leo Breiman and machine learning but my personal experience suggests that if the data satisfies all the assumptions of statistical modelling and we believe in the principle of parsimony, then simple statistical models are better for interpretation. However, if we are interested in classification or prediction then machine learning techniques outperform statistical modelling. This is quite often the case in real life when there is big data with several variables, the assumptions are not satisfied, data are missing or if there are interactions amongst the variables. We can also bridge the gap between statistical models and machine learning by creating a hybrid or ensemble of the models. Raper (2020) has created another debate on the culture of complacency between Professor Cox and Professor Breiman. Should statistical or biometric research start with the hypothesis and then collect and analyse the data or should we start with the data? Unfortunately, in the era of big data and machine learning we start with data which has already been collected.

Coming back to the debate on inductive and deductive reasoning in the broad sense can we consider machine learning (algorithmic based) approach as deductive and statistical modelling as inductive?

Breiman, Leo. Statistical Modelling: The Two Cultures (with comments and a rejoinder by the author). *Statist. Sci* 16 (2001), no. 3, 199--231.

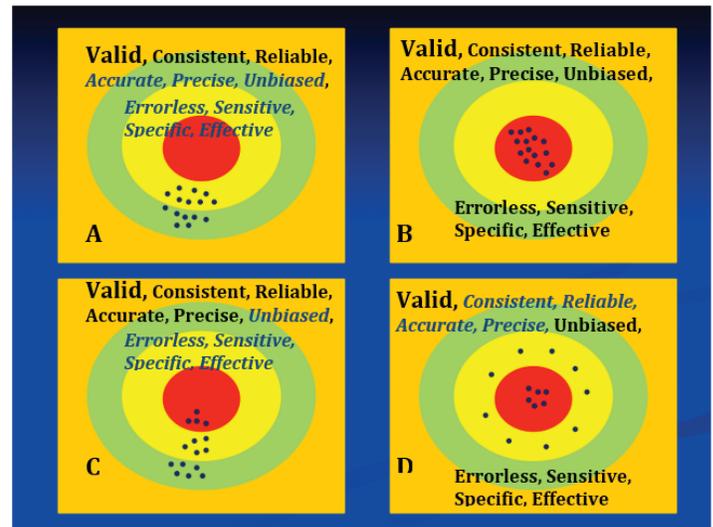
Raper, Simon (2020) Turning points Leo Breiman's "Two Cultures", *Significance*, Volume 17, issue 1, 34-37.

# Basic Themes V & VI (Reader's responses are invited)

**Validity** (Statistical): One of the foremost concepts invariably used in scientific endeavors is validity of the instruments or scales or questionnaires or interview schedules etc.; frequently deployed for evaluation studies or to examine the applicability of tests especially in the fields of psychology, sociology and education. Similarly, essential in teaching & practicing research-methodology are the concepts of validity of tools, techniques, methods (*including design and sampling*) & procedures or even to say the skills to be used to achieve the specific objectives of that focused research. The simple meaning of the word valid derived from the Latin term '*validus*' is 'strong' or 'worthy'. But various disciplines in science have attempted to decode it further in their own ways or to define the term 'strong' or 'valid' by deriving a wide-spectrum concepts such as; *Internal & external validity; face & predictive validity; criterion & content validity; construct & concurrent validity; ecological & population validity;* and so on so forth. However, our focus should rather be limited to understanding the fundamental concepts of validity involved in biometrics that frequently deploy a variety of tools of measurements in experimentations for scientific observations or for validating and evaluating the effect of interventions or actions. Once these concepts are properly understood should easily be applied to a variety of situations encountered in science of research.

We shall agree to the fact that with an objective to measure the weight of an individual subject or object the use of a BP instrument is not at all valid instead the use of a weighing machine would certainly be valid. But we may require simultaneously weighing machines of varied precision to weigh an elephant or an adult human being or to monitor the growth of human fetus or a neonate and similarly still higher precision may be needed for weighing jewels of silver or gold or diamond or platinum and still highly sophisticated and precise equipment is a must to know the weight of an atom and so on. Therefore, it is quite clear that a weighing machine with a fixed calibration or limited precision can never be universal or so versatile to serve a variety of weighing objectives. The above situations do make it clear that even for measuring weight there cannot be something like an 'absolute validity' concept. Therefore, the concept of 'objective' validity should, although be agreed, its applications in performing a variety of specific tasks would certainly invite 'subjectivity' under compulsions of varied understanding. This is bound to happen, as the generous definition of validity suggests, any specific performance (of a tool or definition) to be valid should ensure to achieve what it was **intended or purported to achieve**. But finally this crucial decision on its choice of recruitment for a specific task and implied verification of the end result or achievement has to be left to the subjective and discrete wisdom of the investigator.

An old but well circulated and popular example used in explaining the concept of validity is creating a variety of scenarios (**Figure-1**) based on (*usually 5 or 10*) or **15 rounds** of shooting exercise of a targeted "**bull's eye**". That also is used in deriving the understanding of certain other important quality qualifying terms of interest like; *Error & Bias, Consistency, Reliability, Accuracy, Precision and Sensitivity* etc. that are strongly associated with Validity.



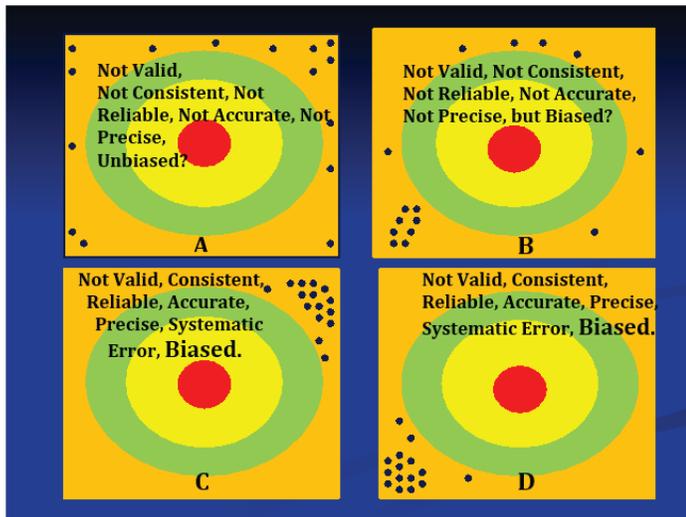
Is there anything wrong if we claim that in all the four Quadrants (Q) - A, B, C & D of Figure -1, each set of fifteen shots conform to the norms of **validity**? Of course, that is depending on the separate criteria defined for each quadrant each time *i.e. either innermost (red) for - Q-B, middle (yellow) for - Q-D or outer circles (green) for - Qs-A & C (the face validity)*. So taking a cue from the above example of using weighing machines of different calibrations the concept of validity has been more focused or narrowed down in Q-B, whereas, less in Q-D, and still loose for Qs-A & C. While we should agree that not even a single quality qualifying terms were compromised or sacrificed in Q-B; Why objectionable become the use of terms like *errorless, unbiased and consistent & reliable; or even accurate & precise*, if these qualities are performing well in their respective zones predefined, as depicted in Q-A, C and D? In fact here should begin the debate on the extent of objectivity and subjectivity involved in decision-making with reference to the prefixed criteria.

In contrast, however, in **Figure -2** in all the four Quadrants, A, B, C, & D, the concept of validity is not at all visible but systematic **errors & bias** are prevailing throughout even for a loose criteria of allowing shooting at least within the outermost green circle, as in each Quadrant all the fifteen shots consistently go beyond the prescribed boundary. Presuming that the recruitments of the shooter as well the shooting equipment were ensured to be valid before the performance, *Is it not a fit case of validity being compromised for the sake of a strong bias?* Is it not happening with ongoing advanced research? Mostly, designing of strict research protocols are ensured in the beginning but frequently get loose or diluted during their implementation part to lead a risk for false interpretations of results and sometimes even to wrong conclusions? While Figure-2 demonstrates an extreme example of violation of validity for easy explanations and understanding, we may notice that most of the researchers choose to construct mixed models with borderline situations to plead their own bias either by mistake or sometimes other way round.

## Basic Theme - VI (Reader's responses are invited)

**Errors and Bias:** Is it a true statement that no experimentation or observation or measurement can either ever be totally free from errors or can always escape from being called biased? Yes, it may probably proved to be true and realized, if and only if the

proper scrutiny to discover that error or bias, is subjected to high order sensitivity. In other words the undetected errors, if any, in question would certainly be identified depending upon the sensitivity and precision of the scrutiny tool. Therefore, it is warranted that the suitability of choosing a microscopic or a Nano-metric or a Pico-metric sensitive scale should be predetermined before its deployment for a critical scrutiny. However, to define it still more clearly, we must understand that even if the observations or measurements are valid, but they do show any lack in desired precision and accuracy, irrespective of the magnitude or quantity or the direction of deviation from the intended measurement should be called errors. Furthermore, one sided repeated **errors** or systematic errors are called **bias**.



In accordance with the above definitions, we conclude that bias is taking place for sure in Quadrants C & D of Figure-2, along with the quality qualifying parameters of consistency and reliability. But we also observe that shooting is once again biased in Quadrants A & B, as all the shots go systematically beyond the outermost green circle; not even a single shot is within red, yellow or green zones. While in later cases the consistency and reliability is lost as the shots are scattered, why not accept the fact that an overall strong bias is in favor of shooting beyond the outermost green circle prevailing in all the four Quadrants, even at the cost of compromising validity? If this overall peculiar shooting is just to be attributed to the error term and not as strongly biased, at least a few shots should have hit somewhere else touching the inner areas of the three circles demanding varied precision and accuracy. This happening repeatedly and deliberately not hitting the targets is a more serious bias than what we have been looking for. It is certainly not the case of hitting a target by a gold-smith versus iron-smith's hammer or crude versus refined actions. Also, it is not the gap between communication and understanding. Rather sometimes it may be deliberate and purposeful for claiming the same is being delivered with entirely a different but biased objective. Capturing these motivated errors and biases or better to call them '**noise**', performed intelligently beyond the set objectives is also equally important considering their consistent prevalence within a large but specific domain; whether virtual or real. Do we notice it happening with everyday science around us? Is it an intelligent intellectual mistake or brainwash? Although many of us do claim to be pure scientists respecting ethical research, after all we are human beings living with our own inherent biases that are 'but natural' and bound to be reflected in our research outputs as well.

Now let us come out of the "bull's eye" scenarios created by us and try to examine an overall bias sometimes sitting in our subconscious mind that cannot be helped out. We understand well that the science of religion and faith is not always to match with principles of modern science, mainly because the two are independently logical and valid but biased to their own domains of philosophical or scientific reasoning. In such a case both bring undesired 'noise' in the functioning of each other. Similarly, the pharmaceuticals drugs practiced in Allopathic medical sciences in contrast to the treating preparations prescribed with their own spiritual and psychological touch in other systems of medicines like AYUSH (*Ayurveda, Yoga, Unani, Siddha and Homeopathy*), and Naturopathy including several other indigenous systems of healing with faith may once again be scientific, logically valid, but biased to their own independent domains of wisdom. None of these systems may be error free or perfect but the comparisons of the bias of any two systems may neither be valid nor feasible at all. Usually the basic philosophy of one system is not understood by the other and therefore, working with divergent distancing is preferred considering that they cannot be complementary to each other. Only a few systems respect each other having something in common and prefer to work together in convergence. The above discussions are summarized to caution that each research process or scientific thinking has the right to claim its own validity and to have its own bias. These overall generalized biases and the question of their validity prevailing around us should be differentiated from the bias and the validity of the tools used in biometrics to deal with specific scenarios satisfying well our logic and reasoning compartments. Realizing there exists a very thin difference between them in practice, extra care is needed into identifying and differentiating their simultaneous but silent happening that frequently is likely to go unnoticed.

We experience these generalized as well specific errors and biases while preparing and implementing research protocols, collecting data, analyzing and transforming data into results, struggling with inferential statistics to derive meaningful conclusions, talking about good estimators in statistics, and so on. There exist **hundreds of errors and biases** specific to statistics in research, design & sampling, interpretation & decision-making, not only limited to 'epidemiology', 'medicine & health', rather more frequently used in 'Agriculture & Forestry', 'Ecology & Environment' and 'Biology & Genetics', as well for a variety of applied research and for many more other scientific activities. However, our objective here is not to enlist a series of errors and biases frequently encountered by the science of biometrics. But we must identify and recognize them for their elimination as far as possible or to control and minimize their effects. For example, selection or allocation biases, measurement biases, instrumental errors & biases, inter & intra investigator or observer's biases, misclassification biases and so on so forth, are some of the frequently encountered biases. We know that the techniques of **blinding, randomization, replication, standardization** of techniques and procedures, selection of appropriate **controls** and to a great extent the **experimental/study designs** do help us to overcome some of them.

# President's Corner

*Continued from p. 1*

able and willing to attend. Second, it is not clear that all the programming planned for July 2020 would automatically roll over to August 2021. All speakers would have to be contacted to check if they would still be willing to keep their presentations, and appropriate replacements would have to be selected for those no longer interested in participating in 2021. This could entail quite a lot of work for the Program Committee and our Business Office which, by then, would already be heavily involved in setting up the IBC2022 meeting.

In summary, the Officers and the Executive Board have decided that the risks associated with postponing IBC to August 2021 are too high for our Society to take and IBC2020 had to be cancelled. Obviously, this affects many IBS members, not in the least the members of the Local Organizing Committee and the International Program Committee who worked very hard, for several years, to make this meeting a success. On behalf of all Officers, I would like to thank them for their efforts. I have received messages from members from all over the world expressing deep sympathy with the organizers. The Officers, the Local Organizing Committee, the Korean Regional Officers, and some Officers from neighbouring regions are currently exploring the idea of having a large multi-region meeting in 2021 which would incorporate some aspects of the original IBC2020 meeting. Of course, we all realize that this can only partially compensate for the cancellation of IBC2020.

In parallel, Louise Ryan, organizing president for the original IBC2020 meeting is exploring the possibility of offering some parts of the original program online and to make those available to our members. For example, Geert Molenberghs, Executive Editor of Biometrics, is currently organizing an online Biometrics Showcase Session during which the best Biometrics papers for 2018 and 2019 will be presented. Brian Reich, our JABES editor, is doing the same. Louise is also exploring having some of the keynote/invited talks take part in the online program. For example, Peter Diggle already agreed to deliver his Presidential Invited Address virtually. Also, the Caucus for Women in Statistics (CWS) has established an award to honour Florence Nightingale (1820-1910) who was an inspirational person with a passion to serve and care, but also was an acclaimed and compassionate statistician. The selected finalists in the Award Competition will present their work during a virtual session and they will be judged by members of the IBS Award Committee. The winner will be announced digitally and recorded versions of the finalists presentations will be made available on the CWS and IBS websites. The entire virtual program will start in July and will be offered over a couple of months. Sessions will be given at a variety of different times to accommodate different time zones and recordings will be made available afterwards. While I fully realize that this online program will never make up for the loss of our IBC2020 meeting, I am confident that it will eventually turn out a good alternative. The program already looks very promising and Louise will continue to extend it. I very much recommend checking out our website ([www.biometricsociety.org](http://www.biometricsociety.org)) where details will be posted as they become available, and I hope many of you will be able to attend several of the virtual sessions.

Speaking of the website, many of you may have noticed that our new website has been launched recently. All IBS functionalities are

ready, and we also have a fully functional IBS Members Community to post questions, to review content from the community, to access materials specifically archived for our members, etc. I very much invite all of you to join the community by signing in to [biometricsociety.org](http://biometricsociety.org) with your existing member credentials, to update your personal information, and to join the Community, and to post questions or topics. This channel is the perfect opportunity to interact with thousands of colleagues worldwide. While the website already incorporates all IBS functionalities, we intend to further extend it in the near future. For example, regional mini-sites will be added as a service to our regions who can use their own mini-site for regional activities and communication.

I already briefly mentioned IBC2022 to be held in Riga (Latvia), July 10-15, 2022. While still a long time ahead of us, we have started initial planning. The conference venue will be the Radisson Blu Latvija Conference & Spa Hotel. Kerrie Mengersen (Queensland University of Technology, Australia) will serve as International Program Chair and she is currently composing the International Program Committee. Once the committee has been approved by the Executive Board, we will start planning the scientific program. I invite you all to start reflecting on interesting proposals for short courses and invited sessions in order to be ready by the time a general call will be sent out.

On a more governance-related note, I am pleased to announce that José Pinheiro, Global Head of Statistical Modeling & Methodology at Janssen Pharmaceuticals, has been selected to serve the Society as President-Elect from 1 January 2021 through 31 December 2021. He will automatically ascend to the office of President for two years, followed by a term of one year as Outgoing President. I very much look forward to working with him. José will need little introduction to our Society because he has been an active member since long. As Louise did two years ago when I was elected, I intend to introduce him as soon as possible into all the governmental details, such that he is fully informed by the time he will take office. I also would like to take the opportunity to thank all members who participated in the online election.

To conclude, I hope I will get to meet many of you at any of our virtual meetings this summer, and I wish you all a good holiday period. For many, the summer period will be different from what originally envisioned, due to the pandemic. Still, I hope you will enjoy a well-deserved break.

**Geert Verbeke**  
**President, IBS**

# International Biometric Conference | Virtual Learning Series



We are delighted to invite you to register to participate in the virtual IBC2020. An exciting menu of offerings is being planned, including keynote addresses by Peter Diggle, Di Cook and David Donoho. Most of the planned invited sessions, our Statistics in Practice ses-

sion, as well as our Young Statistician Showcase, our Biometrics Showcase and our JABES showcase will be offered. We are in the process of engaging with those who were accepted for contributed oral and poster presentations to enable them to participate as well. We are particularly excited about the platform that will be used for the contributed sessions since it allows quite a lot of flexibility in what can be uploaded. As always happens at the in-person IBC, there will be awards for the best poster (Student and Professional category) as well as for best oral presentation by a student.

While the precise timetable is still being worked out, the plan is to offer sessions via ZOOM throughout July and August. The time of the day will vary, depending on the timezone of participants. We are aiming to arrange the timing so that everyone will have the opportunity to participate in at least some sessions live at reasonable times. You will always have the option to watch recordings later. Our very first session will be a Virtual Conference opening session, with President Geert Verbeke introducing Professor Peter Diggle who will deliver the Presidential Invited Address. This session is being planned for 09:00 London time on Monday 6 July 2020. Please mark your calendar! Once you register, you will be put in the system to receive updates as plans progress and the schedule is further locked in.

As you will see when you go to register, we have kept the price very modest. But, IBS members will be able to request a fee waiver if they can not pay. If you can pay, that would be much appreciated since it helps cover costs and also helps recoup some of the losses associated with cancelling the in-person meeting.

We look forward to seeing you online!

**Louise Ryan, IBC2020 Organizing President**  
Renato Assunção, International Program Chair, IBC2020

## Registration Fees:

By registering you will receive full access to all live sessions over multiple weeks, dates & time in July and August of 2020. Registration cost will also go towards software license to run the series, plus a very small amount to help the Society recover cancellation fees from event bookings in Seoul, Korea.

IBS Member	\$100 USD
Non-Member	\$130 USD
IBS Student Member	\$0
Non-Member Student	\$0
Member LMCI	\$0
Non-Member LMIC	\$0
Financial Hardship Request	\$0

*The Financial Hardship option is for IBS Members only - We understand these times are tough and not everyone will be in a position to pay. The standard fees are suggested and will be very helpful to IBS in recovering some of our losses due to the cancellation of the IBC in-person conference. But if you can not afford it, please feel free to pay what you can, even if that is zero. We definitely want you to participate in the conference! Your registration will show as \$0 but you will have the opportunity to add a "Pay What You Can" donation to support the IBC prior to checkout.*

## Awards Update

In keeping with tradition, in even-numbered years, Best Papers in Biometrics and JABES by IBS members awards are identified, for the immediately preceding years. The Co-editors for each journal have selected the following papers for 2018 & 2019.

### Best Paper 2018-2019 - Biometrics

John D. Rice, Robert L. Strawderman, and Brent A. Johnson (2018). **Regularity of a renewal process estimated from binary data.** [Biometrics 74, 566-574](#)

Andrew G. Chapple and Peter F. Thall (2019). **A hybrid Phase I-II/III clinical trial design allowing close re-optimization in Phase III.** [Biometrics 85, 371-381](#).

Normally, the award winning papers are presented in a "Biometrics Showcase" session during the International Biometric Conference. Of course, due to COVID-19, 2020 is a very odd even-numbered year and will not take place in the planned format. We invite you to register for the Virtual IBC2020 and join the online showcase sessions featuring the following papers.

**Session I:** Andrew G Chapple & Peter F. Thall: A hybrid Phase I-II/III clinical trial design allowing dose re-optimization in Phase III

### Thursday, 8 July 2020

- 08:00 Central Time (Chicago)
- 09:00 Eastern Time (New York)
- 14:00 Greenwich Mean Time (London)
- 15:00 Central European Time (Paris, Berlin)
- 23:00 Australian Eastern Standard Time (Sydney)

**Session 2:** John D. Rice, Robert L. Strawderman, and Brent A. Johnson: Regularity of a renewal process estimated from binary data

## Tuesday, 18 August 2020

10:00 Mountain Time (Colorado)

12:00 Eastern Time (New York)

17:00 Greenwich Mean Time (London)

18:00 Central European Time (Paris, Berlin)

02:00 (August 19) Australian Eastern Standard Time (Sydney)

## Best Paper 2018-2019 - JABES

Daniel W. Gladish, Daniel E. Pagendam, Luk J. M. Peeters, Petra M. Kuhnert & Jai Vaze (2018). **Emulation Engines: Choice and Qualification of Uncertainty for complex Hydrological Models.** *Journal of Agricultural, Biological and Environmental Statistics* **23**, 39-62.

Erin M. Schliep, Alan E. Gelfand, James S. Clark, and Roland Kays (2018). **Joint Temporal Point Pattern Models for Proximate Species Occurrence in a Fixed Area Using Camera Trap Data.** [Journal of Agricultural, Biological and Environmental Statistics](#) **23**, 334-357.

Matthew J. Heaton, Abhirup Datta, Andrew O. Finley, Reinhard Furrer, Joseph Guinness, Rajarshi Guhaniyogi, Florian Gerber, Robert B. Gramacy, Dorit Hammerling, Matthias Katzfuss, Finn Lindgren, Douglas W. Nychka, Furong Sun & Andrew Zammit-Mangion (2019). **A Case Study Competition Among Methods for Analyzing Large Spatial Data.** [Journal of Agricultural, Biological and Environmental Statistics](#) **24**, 398-425.

Don't miss out on this session! Register today to receive updates on the JABES Showcase schedule and other Virtual IBC offerings.



## Editorial Update

### Biometrics

#### September 2020 Issue Highlights

The September issue contains a broad area of Biometric Methodology papers, including a discussion paper: "Horseshoe-based Bayesian Nonparametric Estimation of Effective Population Size Trajectories," by James R. Faulkner, Andrew F. Magee, Beth Shapiro, and Vladimir N. Minin. The paper is discussed by Lorenzo Cappello, Swarnadip Ghosh, and Julia A. Palacios. Further papers include "One-step targeted maximum likelihood estimation for time-to-event outcomes," by Weixin Cai and Mark J. van der Laan, "Estimating Individualized Treatment Regimes from Crossover Designs," by Crystal T. Nguyen, Daniel J. Lockett, Anna R. Kahkoska, Grace E. Shearrer, Donna Spruijt-Metz, Jaimie N. Davis, and Michael R. Kosorok, "Order-restricted inference for clustered ROC data with application to fingerprint match-

ing accuracy," by Wei Zhang, Larry L. Tang, Qizhai Li, Aiyi Liu, and Mei-Ling Ting Lee, "Bayesian shrinkage estimation of high dimensional causal mediation effects in omics studies," by Yanyi Song, Xiang Zhou, Min Zhang, Wei Zhao, Yongmei Liu, Sharon L. R. Kardina, Ana V. Diez Roux, Belinda L. Needham, Jennifer A. Smith, and Bhramar Mukherjee, "Optimal approximate conversions of odds ratios and hazard ratios to risk ratios," by Tyler J. VanderWeele, and "Learning-based biomarker-assisted rules for optimized clinical benefit under a risk-constraint," by Yanqing Wang, Ying-Qi Zhao, and Yingye Zheng.

Also the Biometric Practice section includes a discussion paper: "Testing small study effects in multivariate meta-analysis," by Chuan Hong, Georgia Salanti, Sally Morton, Richard Riley, Haitao Chu, Stephen E. Kimmel, and Yong Chen, and discussed by Hans C. van Houwelingen, by Hisashi Noma, and by James R. Carpenter, Gerta Rucker, and Guido Schwarzer. The section further boasts articles on "Dynamic modelling of multivariate dimensions and their temporal relationships using latent processes: application to Alzheimer's disease," by Bachirou O. Tadde, Helene Jacqmin-Gadda, Jean Francois Dartigues, Daniel Commenges, and Cecile Proust-Lima, "Joint analysis of single cell and bulk tissue sequencing data to infer intra-tumor heterogeneity," by Wei Sun, Chong Jin, Jonathan A. Gelfond, Ming-Hui Chen, and Joseph G. Ibrahim, and "Sample size and power for the weighted log-rank test and Kaplan-Meier based tests with allowance for non-proportional hazards," by Godwin Yung and Yi Liu.

#### Best Papers in Biometrics Awards for 2018 and 2019 announced

In keeping with tradition, in IBC years, Best Papers in Biometrics by IBS members awards are identified, for the immediately preceding years. The Co-editors selected:

John D. Rice, Robert L. Strawderman, and Brent A. Johnson (2018). Regularity of a renewal process estimated from binary data. *Biometrics* **74**, 566-574.

Andrew G. Chapple and Peter F. Thall (2019). A hybrid Phase I-II/III clinical trial design allowing dose re-optimization in Phase III. *Biometrics* **85**, 371-381.

Normally, the award winning papers are presented in a "Biometrics Showcase" session during the International Biometric Conference. Of course, due to corona, 2020 is a very odd even-numbered year. Thus, because of IBC's cancellation, the session will not take place but the awardees will nevertheless be presented with their award.

#### 2021 - 2023 Co-editor for Biometrics identified

The search committee to identify a successor for Co-Editor Debashis Ghosh, whose term will end 31 December 2020, has identified William F. (Bill) Rosenberger (George Mason University, Virginia, US) as Debashis's successor. Bill will serve from 1 January 2021 to 31 December 2023. The search committee consisted of Geert Molenberghs, Biometrics Executive Editor, Chair (Belgian Region); Debashis Ghosh, Biometrics CE (ENAR); Mark Brewer, Biometrics CE (British & Irish Region); Alan Welsh, Biometrics CE (Australasian Region), Hans-Peter Piepho, EAC Chair (German Region); Rebecca Hubbard, EAC member (ENAR); Andrea Lavalle, EAC member (Argentinian Region).

## Associate Editor Panel turnover

The panel of Associate Editors is composed of colleagues who each serve two-year terms. This means that every 1 July a number of them end their terms. The Co-Editors have invited new Associate Editors to replace colleagues retiring and, in view of the large number of submissions received, also to expand the size of the panel. New Associate Editors are chosen to replace expertise lost by retiring Associate Editors, as well as to fill in gaps in expertise in certain areas, especially in emerging fields. The panel will consist of nearly 100 members, starting 1 July 2020.

We are grateful for the service, rendered by our retiring Associate Editors: Hongyuan Cao, Peter Gilbert, Jelle Goeman, Yi Li, Wenbin Lu, Yanyuan Ma, Micha Mandel, Gary Rosner, Rui Song, Lu Tian, Yingqi Zhao, and Cory Zigler.

## Biometrics' Response and Resilience Towards COVID-19

At a time where we all explore uncharted territory, Biometrics has been receiving a higher than usual number of submissions since the month of February. We are of course delighted by this trend. While both quality and timeliness of the peer review process is always at the fore, we prioritize to the extent possible papers with a COVID-19 theme. Our large and further increased Associate Editors panel, comprised of very dedicated colleagues, ensures qualitative review even in these difficult times. In our communication with Associate Editors and referees, we routinely acknowledge the potential for disrupted schedules due to the pandemic. The current three Co-editors, Debashis Ghosh, Mark Brewer, and Alan Welsh, can count on additional editorial support from their immediate predecessors, Stijn Vansteelandt and Malka Gorfine, should health- or work-related obstacles interfere with their work. Also Marie Davidian, former Executive Editor, is at the ready to provide help if needed.

Because of a severe reduction in passenger flights worldwide, and the fact that precedence needs to be given to the transportation of medical and other vital materials, it is currently not possible to ship paper copies of Biometrics, meaning that the publication is temporarily electronic only. Articles can be found, soon after their acceptance, on Wiley's site, accessible directly or via the IBS's web pages. As soon as the situation allows, regular paper copy shipping will be resumed. Wiley is doing the needful to ensure that their Biometrics workflow is protected, both on site as well as for the work done by vendors.

**Geert Molenberghs, Executive Editor Biometrics**

## JABES

This spring we are very excited to announce the top papers that have appeared in JABES from 2018-2019. The winning papers are:

“Emulation Engines: Choice and Quantification of Uncertainty for Complex Hydrological Models” by Daniel Gladish, Daniel Pagendam, Luk Peeters, Petra Kuhnert and Jai Vaze.

“Joint Temporal Point Pattern Models for Proximate Species Occurrence in a Fixed Area Using Camera Trap Data” by

Erin Schliep, Alan Gelfand, James Clark and Roland Kays.

“A Case Study Competition Among Methods for Analyzing Large Spatial Data” by Matthew J. Heaton et al.

These papers make great advancements to environmental statistics and will be presented at the next International Biometric Conference. Congratulations to the authors and thank you for your excellent contributions!

We are also excited to announce the June issue. The issue is comprised of the following papers: “Estimating Reproduction and Survival of Unmarked Juveniles Using Aerial Images and Marked Adults” by Perry Williams, Cody Schroeder and Pat Jackson; “Statistical Development of Animal Density Estimation Using Random Encounter Modelling” by N. O. A. S. Jourdain, D. J. Cole, M. S. Ridout and J. Marcus Rowcliffe; “Sampling Strategies to Estimate Deer Density by Drive Counts” by Lorenzo Fattorini, Alberto Meriggi, Enrico Merli and Paolo Varuzza; “Modeling Partially Surveyed Point Process Data: Inferring Spatial Point Intensity of Geomagnetic Anomalies” by Kenneth Flagg, Andrew Hoegh and John Borkowski; “Estimation in Complex Sampling Designs Based on Resampling Methods” by Bardia Panahbehagh; “Projecting Flood-Inducing Precipitation with a Bayesian Analogue Model” by Gregory Bopp, Benjamin Shaby, Chris Forest and Alfonso Mejía; and “Obtaining a Balanced Area Sample for the Bureau of Land Management Rangeland Survey” by Cindy Yu, Jie Li, Michael Karl and Todd Krueger.

If you have a suggestion for a special issue, we would be pleased to hear from you. We are also keen to publish papers that summarize the state of methodological development in subject areas for which technological advances are generating a demand for new statistical approaches. If such papers also speculate on likely future developments, so much the better. If you feel that you could offer such a paper, or can suggest a topic together with possible authors, please let me know.

For more information on upcoming issues, the editorial board, and the aim and scope of the journal, please visit our website <http://link.springer.com/journal/13253>. We also accept submissions of books to review in the upcoming issues of JABES; to submit a book for review, please see the above website (click on “Editorial Board”) or contact Eleni Matechou ([e.matechou@kent.ac.uk](mailto:e.matechou@kent.ac.uk)).

Please follow us on Twitter: @JabesEditor.

**Brian Reich, Editor in Chief**

# Software Corner

## Power and Sample Size Calculations by Simulation with R and `simglm`

Brandon LeBeau, [brandon-lebeau@uiowa.edu](mailto:brandon-lebeau@uiowa.edu), University of Iowa

Statistical power is the probability that a statistical analysis is able to detect a non-zero population effect for a binary hypothesis test. Power and sample size calculations are commonly done prior to collecting data. Doing a power analysis that mimics the data to be collected can help to ensure that the sample collected is large enough to detect the effect of interest and also overcome sample data that do not meet the assumptions of the statistical model. To ensure the power analysis is as similar to the collected data as possible, the analyst must have a good grasp of the data to be collected, which would include estimating the size of the effect to be detected, study design, amount of missing data that can be expected, among many others. There are at least two ways to conduct a power analysis, one using closed-form solutions or using Monte Carlo simulation methods.

### Closed-Form Power Analyses

Power is often evaluated using closed form solutions that assume the assumptions for the statistical model are not violated. For example, residuals for many statistical analyses are assumed to follow a normal distribution and are also assumed to have homogeneity of variance. If the data collected do not meet these assumptions, the estimated power from the closed form solutions will differ from the power from the real data, and in many cases are likely to be overestimated in the presence of assumption violations like heterogeneity of variance.

Statistical software has been developed to estimate power when statistical assumptions have been made such as [G\\*Power](#) (Faul et al. 2007), [PowerUp!](#) (Dong et al. 2015), or [Optimal Design](#) (Raudenbush et al. 2011). There are also packages in statistical software programs such as [pwr](#) (Champely 2018), [WebPower](#) (Zhang and Mai 2018), [stats](#) which is a part of base R (R Core Team 2019) or [statsmodels](#) in Python that estimate power for relatively simple statistical analyses such as t-tests, analysis of variance (ANOVA), linear regression, correlation, or general linear models. G\*Power implements power for similar analyses as those found in traditional statistical software implementations, but offers a graphical user interface (GUI) that may aid users in the power estimation. Finally, the specialized software, PowerUp! and Optimal Design are commonly used to estimate power for randomized control trials and when there are nesting effects that are common in educational or psychological research.

### Power by Simulation

Power by simulation differs from the traditional approaches in the degree of flexibility the user has on the power analysis. For example, users can directly specify the generating model, analysis model, and all aspects of the generating and analysis model(s). This flexibility may better align the power analysis with the data collected where statistical assumption violations often occur. In these situations, it would be useful to know the impact on power these types of violations may have, especially if the closed-form solution power estimates are overestimated.

The following are general steps that are taken in a power by simulation example.

1. Assume population parameters, including
  - population effect size(s) of interest.
  - distribution of variable(s) and residuals.
  - variance of variable(s) and residuals.
2. Simulate data based on a statistical model.
3. Fit a statistical model to the simulated data.
4. Replicate steps 1 - 3 many times.
5. Calculate the proportion of statistical tests that reject the hypothesis.

In the power by simulation framework, data are simulated based on assumed values for the population, for example the population mean difference between two groups. Given that these values are true, data are simulated many times and a statistical model is fitted to the simulated data. To estimate power in this framework, the number of statistical tests that reject the binary hypothesis compared to the number of replications provides an empirical estimate of power. The largest drawback to the simulation approach is the time required to run the power analysis. Another drawback is the statistical programming skills required to perform the simulated power analysis is greater than closed-form software that often includes a graphical user interface.

There are R packages to conduct power by simulation, including **clusterPower** (Kleinman, Moyer, and Reich 2017), **longpower** (Donohue 2019), **PAMM** (Martin et al. 2011), **powerlmm** (Magnusson 2018), **simr** (Green and MacLeod 2016), and **simglm** (LeBeau 2019). The focus of the rest of this discussion is around the `simglm` R package for power by simulation.

### Introducing `simglm`

The **simglm** package offers users the ability to simulate data from general(-ized) linear (mixed) models and contains wrappers that aids users in replicating the analysis. Once the data have been simulated, the package contains functionality to fit models to the generated data for power estimation. The most recent version of the package, v 0.7.4, fully implemented a new “tidy” syntax in which functions are pipeable and follow a more consistent syntax, while increasing the flexibility for the elements of the data generation. The package uses the **future.apply R** package for parallel processing to speed up the data generation and power estimation.

The **simglm** R package has the ability to simulate continuous, binary, and count outcome data from the general linear model with and without random effects. The user has the ability to add heterogeneity of variance, various missing data mechanisms including MCAR, MAR, and dropout missing for longitudinal data, and specify different distributional assumptions for fixed and random model attributes. When specifying random effects, these effects can be specified as nested or cross-classified. In addition, the current developmental version has the ability to simulate polynomials with the `poly()` function and will soon be able to simulate splines models from the **splines** package with the `ns()` or `bs()` functions.

When using the **simglm** package for a power analysis, the user has the flexibility to specify their own model for the model fitting step (step 3 mentioned above). This gives the user the flexibility to explore the impact of power when the model is misspecified

or the model fitting procedure differs from the model generating procedure. For example, this could be useful if the model generating procedure was a three-level linear mixed model with 5 clusters at level three. In these situations, a three-level model may not be feasible in the model fitting stage due to fewer level three units, therefore two-level models could be fitted or a generalized estimating equations (GEE) approach could be specified instead to understand how this may impact the power estimate. Users can also specify the alpha value, type of test statistic, and vary terms. For example, users can specify a variety of sample sizes, in which case a factorial design is used for all varied terms and all varied conditions are explored.

### Additional Resources

Additional resources for the `simglm` package can be found on my website, [brandonlebeau.org](http://brandonlebeau.org). In addition, package vignettes can be found on the [CRAN page for the package](https://CRAN.R-project.org/package=simglm), developmental versions and issues can be submitted on the [simglm package GitHub page](https://github.com/brandonlebeau/simglm), and a pkgdown site is [available](https://brandonlebeau.org/simglm/).

### References

Champely, Stéphane. 2018. *Pwr: Basic Functions for Power Analysis*. <https://CRAN.R-project.org/package=pwr>.

Dong, Nianbo, Benjamin Kelcey, Rebecca Maynard, and Jessaca Spybrook. 2015. *PowerUp! Tool for Power Analysis*. <https://www.cau-salevaluation.org/power-analysis.html>.

Donohue, Michael C. 2019. *Longpower: Power and Sample Size Calculations for Linear Mixed Models*.

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. "G\* Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences." *Behavior Research Methods* 39 (2): 175–91.

Green, Peter, and Catriona J. MacLeod. 2016. "Simr: An R Package for Power Analysis of Generalised Linear Mixed Models by Simulation." *Methods in Ecology and Evolution* 7 (4): 493–98. <https://doi.org/10.1111/2041-210X.12504>.

Kleinman, Ken, Jon Moyer, and Nicholas Reich. 2017. *ClusterPower: Power Calculations for Cluster-Randomized and Cluster-Randomized Crossover Trials*. <https://CRAN.R-project.org/package=clusterPower>.

LeBeau, Brandon. 2019. *Simglm: Simulate Models Based on the Generalized Linear Model*. <https://CRAN.R-project.org/package=simglm>.

Magnusson, Kristoffer. 2018. *Powerlmm: Power Analysis for Longitudinal Multilevel Models*. <https://CRAN.R-project.org/package=powerlmm>.

Martin, Julien GA, Daniel H Nussey, Alastair J Wilson, and Denis Reale. 2011. "Measuring Individual Differences in Reaction Norms in Field and Experimental Studies: A Power Analysis of Random Regression Models." *Methods in Ecology and Evolution* 2 (4): 362–74. <http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2010.00084.x/abstract>.

Raudenbush, Stephen, Jessaca Spybrook, Howard Bloom, Richard Congdon, Carolyn Hill, and Andres Martínez. 2011. *Optimal Design Software for Multi-Level and Longitudinal Research*. <http://hlmssoft.net/od/>.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Zhang, Zhiyong, and Yujiao Mai. 2018. *WebPower: Basic and Advanced Statistical Power Analysis*. <https://CRAN.R-project.org/package=WebPower>.

# STRENGTHENING ANALYTICAL THINKING FOR OBSERVATIONAL STUDIES (STRATOS)

## Introducing the Simulation Panel (SP)

Anne-Laure Boulesteix (1), Tim Morris (2), Willi Sauerbrei (3) and Michal Abrahamowicz (4) on behalf of the Simulation Panel

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

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

(3) Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Freiburg, Germany

(4) Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada

Previous issues of this Bulletin introduced the nine Topic Groups (TGs) of the STRATOS initiative.

Soon after the formal launch of the initiative at the 34th annual meeting of the International Society for Clinical Biostatistics (ISCB) in August 2013, we realized the need for several cross-cutting panels. The first panels were created to deal with 'internal' issues such as applications for STRATOS membership (Membership Panel), 'rules' for talks and papers on behalf of the STRATOS initiative or one of its TGs (Publication Panel), and the need for TGs and panels to employ common terminology (Glossary Panel). See the STRATOS newsletter from May 2018 for more details. Meanwhile STRATOS has eleven panels and this BB report is the first of several, each of which will introduce one of them.

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. Consequently, in 2015 the STRATOS steering group decided to start a Simulation Panel (SP).

The two co-chairs of the SP are Michal Abrahamowicz and Anne-Laure Boulesteix. Further members are Harald Binder, Rolf Groenwold, Victor Kipnis, Jessica Myers Franklin, Tim Morris, Willi Sauerbrei, Pamela Shaw, Ewout Steyerberg, and Ingeborg Waernbaum.

The aim of the Simulation Panel is to provide guidance for the design, implementation, interpretation and reporting of simulation studies for different types of audience (from data analysts with limited statistical background to experts). More generally, the panel is interested in methodological aspects of the empirical assessment of the performance of statistical methods in simulated and/or real-life data. One of the basic principles adopted by the initiative is that the guidance to be formulated in documents written by STRATOS TGs is based on solid evidence, including both theoretical considerations and evidence from empirical studies comparing and validating the relevant methods (Sauerbrei et al 2014). Independently of STRATOS, reliable empirical evidence is frequently needed everyday by data analysts facing choices regarding the use of one or more of the alternative existing methods in their analysis. Systematic evidence of the behavior and performance of methods can be obtained through simulations, but how can we ensure such evidence is robust and reliable? This is the general question addressed by the Simulation Panel.

Data analysts, knowingly or unknowingly, choose methods based on recommendations that were often derived from simulation studies. Yet, for researchers with limited statistical background interpreting published simulation studies is challenging, not to speak of running their own ones. On the other hand, methodological statisticians frequently need to conduct complex simulation studies to assess the methods under investigation in realistic settings. However, they may face a lack of guidance on the design, implementation and reporting of such studies. A recent tutorial paper offers first guidance for design, execution, analysis, reporting, and presentation (Morris et al 2019). The authors provide a structured approach for planning and reporting simulation studies, which involves defining aims, data-generating mechanisms, estimands, methods, and performance measures (“ADEMP”).

The Simulation Panel published a letter to the Editors of *Biometrical Journal* 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, such as those intended by STRATOS Topic Groups, which assess methods without intention to show the superiority of a particular method; see also Boulesteix et al. (2017) for more insights on the concept of neutral study. 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. Ideas related to the content of this letter and more general issues of simulation studies have been presented at various conferences including the CEN-ISBS in Vienna (2017) and DAGSTAT in Munich (2019). Another talk was accepted for the invited session at IBC 30 in Seoul in 2020.

The panel has a forthcoming paper on simulation studies aimed at a level I audience (Boulesteix et al, 2020). The paper offers a gentle introduction to simulation studies for data analysts and researchers who have little or no hands-on experience in this area but (i) may rely on previously published simulation studies to choose their statistical methods and/or (ii) wish to perform own simulation studies and need to understand the basic principles of designing and conducting such studies.

Led by Tim Morris and Willi Sauerbrei, several SP members have started a project to extend the ADEMP structure from Morris et al (2019) to a ‘profile’ to improve and standardise the reporting in simulation studies, with an emphasis on analyses of simulated

datasets. Basic ideas are shown in Table I of De Bin et al (2020). Further, several members of SP are also involved in a paper of TG9 on simulation studies for high-dimensional data.

The Simulation Panel regards simulation studies as a useful ‘generic’ methodological tool for a broad range of researchers, and attempts to promote their more frequent and more accurate use. On the one hand, applied analysts could be empowered by better understanding the concepts regarding interpretation of simulation results. On the other hand, methodological researchers would do well to pay better attention to neutral and clinically plausible design, sensitivity analyses with variation of crucial assumptions and parameters, and clear reporting. We aim to target both with the aim of enabling researchers to make more informed decisions about the methods they use, based on better evidence.

## References:

- De Bin R, Boulesteix AL, Benner A, Becker N, Sauerbrei W (2020). Combining clinical and molecular data in regression prediction models: insights from a simulation study. *Briefings in Bioinformatics* bbz136.
- 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, Groenewold R, Abrahamowicz M, Binder H, Briel M, Hornung R, Morris T, Rahnenführer J, Sauerbrei W (2020). An introduction to statistical simulations in health research (submitted).
- 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 Methodology* 17:138.
- Morris T, White I, M Crowther M (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 STRATOS initiative (2014). *Strengthening Analytical Thinking for Observational Studies: The STRATOS initiative*. *Statistics in Medicine* 33: 5413-5432.

# Region News

## Eastern North American Region (ENAR)

<https://www.enar.org/>

### WebENARs

Be sure to check the ENAR Webinar website for updates regarding the upcoming WebENAR series, as well as for links to past WebENARs and their recordings: <http://www.enar.org/education/index.cfm>.

### 2020 ENAR Spring Meeting, 22-25 March

The 2020 ENAR Spring Meeting was held 22-25 March as a virtual conference due to the COVID-19 pandemic. In just a few short weeks, program organizers canceled the in-person meeting and converted the conference onto a virtual platform. Still, over 700 participants enjoyed roughly 500 abstracts presented across 108 sessions. Five short courses and five tutorials were also presented virtually. The Sunday evening poster session, roundtables, and the Fostering Diversity in Biostatistics Workshop were not held this year.

ENAR extends congratulations to John Van Ryzin Award winner Zhe Sun of the University of Connecticut and all recipients of the Distinguished Student Paper Award: Rebecca Anthopolos, Columbia University; Teng Fei, Emory University; Lin Dong, North Carolina State University; Guillermo Granados, King Abdullah University of Science & Technology; Yichen Jai, University of Pittsburgh; Yunchuan Kong, Emory University; Yan Li, University of Connecticut; Wodan Ling, Columbia University; Peng Liu, University of Pittsburgh; Yusha Liu, Rice University; Lan Luo, University of Wisconsin-Madison; Xin Ma, Emory University; Stefani Mokalled, Clemson University; Arman Oganisian, University of Pennsylvania; Chan Park, University of Wisconsin-Madison; Bo Wei, Emory University; Guanyu Yang, University of Michigan; Bingxin Zhao, University of North Carolina Chapel Hill; Jincheng Zhou, University of Minnesota.

### 2020 JSM, 1-6 August

The 2020 Joint Statistical Meetings will also be held as a virtual meeting from 1-6 August due to uncertainty surrounding the COVID-19 pandemic. The theme of the 2020 meeting is "Everyone Counts: Data for the Public Good." ENAR has been instrumental in helping to put together an outstanding program and thanks everyone who put forth one of the many proposals for invited and topic-contributed sessions. The 12 invited sessions primarily sponsored by ENAR will cover topics about oncology drug development, semiparametric inference, methods for precision medicine, genetic/genomic and microbiome data analysis, causal inference with interference, graphical models, electronic health record and public data analysis, and mentoring. The four topic-contributed and two contributed sessions include presentations on spatial, functional, and neuroimaging data, informative cluster size methods, network-based studies and meta-analyses, rare disease drug development, and Bayesian methods. The JSM online program can be accessed at <http://www2.amstat.org/meetings/jsm/2020/program.cfm>. ENAR extends a huge thank you to Jeremy Gaskins of University of Louisville for serving on the Program Committee for the 2020 JSM.

### 2021 ENAR Spring Meeting, 14-17 March, Baltimore, MD, USA

The 2021 ENAR Spring Meeting is scheduled to take place in-person in Baltimore, Maryland at the Baltimore Marriott Waterfront. Formal invited paper session, short course, and tutorial course suggestions are under review by the Program Committee. To informally suggest ideas, topics or names of potential speakers, contact Program Chair Howard Chang ([howard.chang@emory.edu](mailto:howard.chang@emory.edu)) or Associate Program Chair Yize Zhao ([yize.zhao@yale.edu](mailto:yize.zhao@yale.edu)). Questions about short course and tutorial courses should be directed to the Educational Advisory Committee (EAC) Chair Brent Coull ([bcoull@hsph.harvard.edu](mailto:bcoull@hsph.harvard.edu)). The local arrangements chair is Vadim Zippunikov ([vzipunn1@jhu.edu](mailto:vzipunn1@jhu.edu)).

### Future Meetings

Stay tuned for information about future meetings, including the 2021 Joint Statistical Meetings to be held in Seattle, Washington, USA from 1-12 August, 2021 and the 2022 ENAR Spring Meeting to be held in Houston, Texas, USA from 27-30 March, 2022.

## French Region (SFB)

<https://sfb.pages.math.cnrs.fr/asso/>

Some news from the French Region of the IBS "Société Française de Biométrie (SFB)".

### Daniel Schwartz Award

The SFB decided to pay tribute to Daniel Schwartz, a French biometrician who founded the SFB in 1949, by creating the "Daniel Schwartz Award" for the best thesis defended recently in the biometric field. Two referees from the SFB committee members were designated to review each nomination. The award comprises every two years a prize of 1000€. The final decision of the committee was to share this award among two young researchers:

Simon Bussy for his PhD entitled: "Introduction of high-dimensional interpretable machine learning models and their applications" supervised by Pr. Agathe Guilloux and Dr. Anne-Sophie Jannot at Sorbonne Université, Paris in 2019.

Corentin Segalas for his PhD entitled "Inférence dans les modèles à changement de pente aléatoire : application au déclin cognitif pré-démence" (Inference on random changepoint models: application to pre-dementia cognitive decline) supervised by Dr. Jacquin-Gadda at ISPED Bordeaux in 2019.

### Scientific Events

1. A session of the SFB was planned within the annual conference of "the French Statistical Society", at the campus Valrose of the Université de Côte d'Azur, Nice (<http://jds2020.sfd.asso.fr>). The conference is postponed to spring 2021.
2. The young biometric researchers' day "Journée des Jeunes Chercheurs en Biométrie" previously planned for April 2nd at the Cnam (Conservatoire National des Arts et Métiers), Paris is postponed. The proposed contributions can be forwarded to the SFB-GDR days in October, 2020.

3. The [SFB-GDR days](#) organized jointly with the “GDR Statistique et Santé”, a research group in Statistics and Health will be held 1-2 October at CNAM (Paris). We will have the opportunity to proceed to the ceremony of Daniel Schwartz award.
4. In spring 2021, the Channel Network conference will be held 7-9 April at CNAM in Paris. Hope that COVID-19 does not transform to COVID-21.

## Japanese Region (JR)

<http://www.biometrics.gr.jp/>

### The 2020 Annual Meeting of the Biometric Society of Japan

The Biometric Society of Japan (BSJ) canceled the annual meeting scheduled for 15-16 May 2020 at the Katsushika Campus of Tokyo University of Science, Tokyo, Japan, due to the ongoing pandemic.

### The 2020 Japanese Joint Statistical Meeting

The BSJ is one of the six sponsoring organizations of the meeting and the 2020 Japanese Joint Statistical Meeting will be held on 8-12 September at Toyama International Conference Center and Toyama Public Hall in Toyama, Japan. The BSJ is organizing an invited session entitled “Statistical methods that evaluates the treatment effect based on the relationship of whether the outcomes are larger or smaller.” The society is also organizing an invited session in which the winner of the Young Biostatistician Award conferred by the society will present his research.

**Ikuko Funatogawa**

#### Meetings:

8-12 September

The 2020 Japanese Joint Statistical Meeting

Toyama International Conference Center and Toyama Public Hall, Toyama, Japan

## The Netherlands Region (ANed)

<https://www.vvsor.nl/biometrics/>

Last Autumn, 8 November 2019, the BMS-ANed organised a meeting jointly with Rianne Jacobs, the winner of the Hans van Houwelingen award 2018:

### “BIOSTATISTICS MEETS EPIDEMIOLOGY”



The event took place at the spectacular Linnaeusborg Building of the University of Groningen, see the above photo. The title of the program reflected very well the focus of the meeting and the afternoon was well received by both epidemiologists and statisticians, including Hans van Houwelingen himself. In

addition the General Assembly of the BMS (Biometric Section of the Dutch Society for Statistics and Operations Research -VVS) and ANed (Dutch region of the International Biometric Society -IBS) was held.



**Rianne Jacobs**, statistician at the University of Groningen and winner of the Hans van Houwelingen award, started off with the topic of her winning paper: On needles and haystacks: Finding answers in complex health data. Many health-related research questions boil down to some form of variable selection – finding a few variables or groups of variables in a complex dataset that

are highly predictive of possibly multiple outcomes. Using the 2012 Salmonella Thompson outbreak in the Netherlands, she presented a Bayesian variable selection (BVS) methodology to discover which foods from which stores might have caused the food-borne disease outbreak. The BVS methodology incorporates missing value imputation and a misclassification correction and was shown to outperform the frequentist alternatives.

**Nadine Binder**, statistician at the Medical Center University of Freiburg, presented “Bridging the gap: From time-to-event methods to their application in a Framingham Heart study reanalysis”. Using as an example an analysis of the Framingham data which reported a significant linear decline in the incidence of dementia over the last 30-40 years (NEJM 2016), she showed how missing disease information due to death results in bias. In this study, incidence of dementia was determined only at certain follow-up timepoints and was unknown for subjects who died between visits. Censoring these death cases at time of death or at the last visit observed disease-free yields an underestimated dementia incidence, a bias which is likely to increase in magnitude as the cohort ages and death becomes more prevalent. For such data, the illness-death multistate model might be a more suitable method than the Cox proportional hazards model. The presentation was sponsored by the [Aletta Jacobs School of Public Health](#).

**Eva Corpeleijn** – epidemiologist, University Medical Center Groningen, presented “Lifestyle over the life course”, on developments in the field of healthy living and healthy ageing. The importance of a healthy lifestyle, e.g. diet and physical activity, is more and more acknowledged in the prevention of chronic diseases, as people grow older, and care costs are rising. Most of our care is focused at the one cause – one cure principle. As this no longer holds in an ageing population, the movement in ‘lifestyle medicine’ is growing and asks for new complex methods.

**Saskia le Cessie** – statistician, Leiden University Medical Center, presented “On causal questions and statistical answers”, based on work performed by the causal inference topic group (TG7) of the international initiative of Strengthening Analytical Thinking for Observational Studies (STRATOS). Randomisation fails as an instrument for many post randomisation exposures. Using an example of a breastfeeding encouraging program, she discussed the different estimands and results and presented a simulation learner.

Abstracts and slides can be found at:

<https://www.vvsor.nl/biometrics/events/biostatistics-meets-epidemiology/>.

## Spanish Region (SEB)

<http://www.biometricsociety.net/>

### 200 YEARS OF FLORENCE NIGHTINGALE

To commemorate the 200th anniversary of the birth of [Florence Nightingale](#), pioneer of modern statistics and epidemiology, and the first woman admitted to the [Royal Statistical Society](#), the [Rladies](#) Spain organization (nodes of Barcelona, Madrid, and Bilbao), in collaboration with the [Spanish Biometric Society](#) and the [Catalan Statistical Society](#) are pleased to announce the data analysis competition: “**200 YEARS OF FLORENCE NIGHTINGALE**”.

The competition aims to analyze the data set that Florence Nightingale used to determine the causes of death of the British Army during the Crimean War, and which served to analyze the factors associated with high mortality and to reduce it significantly.

Further information as well as the Guidelines for participation are available from the website of the [Spanish Biometric Society](#).



*Florence Nightingale (May 12, 1820 - August 13, 1910)*

### Signature of collaboration agreements.

The Spanish Biometric Society (SEB) has just signed collaboration agreements both with the Spanish Plant Pathology Society (SEF) and the Spanish Society of Statistics and Operations Research (SEIO). These agreements will make it possible to strengthen the alliances between SEB and each of these scientific societies by organizing joint conferences and disseminating their activities. In particular, SEB and SEF share common areas of work such as population analysis and disease epidemiology and this agreement aims to promote and facilitate synergies between plant pathologists and biostatisticians. Also, SEB and SEIO share their common interest in statistical methods which can be developed further under this agreement.



## Western North American Region (WNAR)

<http://wnar.org/>

### 2020 WNAR/IMS/KISS/JR meeting

Due to the health concerns associated with the ongoing pandemic and complying with the government's guidance, WNAR decided to cancel the 2020 meeting in Anchorage, Alaska (June 14 - 17,

2020). WNAR is planning to hold its 2021 Annual Conference in Anchorage, Alaska. Dates and venues for the 2021 meeting are still being finalized.

Registration information and other details about the 2021 meeting will be on the WNAR web page [www.wnar.org](http://www.wnar.org) as they become available. We note that WNAR supports speakers from lower and middle income countries to attend the WNAR meeting; the policy statement can be found here.

We extend a huge thank you to everyone who participated in planning the 2020 meeting, including the many researchers who contributed abstracts to the meeting. In particular, we thank our Local Organizer Jiaqi Huang and Program Chair Yingqi Zhao, both of whom have generously agreed to continue in those positions for the 2021 meeting. We also want to thank the Student Paper Committee Chair, Harold Bae, and all the members of that committee for continuing the Student Paper Competition, which will be held virtually in June 2020.

**Megan Othus**

## Announcements & Upcoming Events

### ASA Sponsors New Health Tech Podcast



Looking to stay up to date on developments in health care technology around the world? The American Statistical Association is sponsoring “The Pod of Asclepius”, a new podcast where data scientists, statisticians, engineers, and regulatory experts discuss the technical challenges in their healthcare domain.

We have over 30 episodes published and available on YouTube, Podbean, iTunes, Stitcher, Podchaser, and Tune In Radio.

Looking for a good place to start? Check out the following episode links:

- [Risks and Opportunities of AI in Clinical Drug Development](#) with David Madigan and Demissie Alemayehu
- [Kidney Injury - Biomarkers for Prediction and Prognosis](#) with Allison Meisner
- [Data Platforms to Monitor Animal Health](#) with Shane Burns
- [Bayesian Approaches in Medical Devices: Part 1, Part 2, Part 3](#) with Martin Ho and Greg Maislin
- [What is a Stepped-Wedge Trial?](#) with Mona Kanaan and Ada Keding
- [Machine Learning for Patient Vital-Sign Monitoring](#) with Glen Wright Colopy

You can catch up on all episodes on our YouTube playlist for [Season 0](#) and [Season 1](#).

The easiest way to catch a new episode is to subscribe via our channels ...

**Mail List:** <https://www.podofasclepius.com/mail-list>

**YouTube:** <https://www.youtube.com/channel/UCEz2tDR5K6AjlKw-JrV57w>

**Podbean:** <https://podofasclepius.podbean.com>

We would like to thank the IBS for its assistance in promoting the new series!

## IBS Journal Club

The Journal Club is open to all IBS members free of charge. The primary purpose of the Journal Club, apart from presenting worthy papers in a more public setting, is to widen the scope for understanding these papers and to provide a new networking opportunity for IBS members through a regular internet forum. All sessions are recorded and are available on the IBS website here, <http://www.biometricsociety.org/education/video-sessions/>. To access the recording you must login to your IBS account.

June and August have been postponed due to the Virtual IBC taking place during those months. Join us! [Register today at www.ibc2020.org](http://www.ibc2020.org).

## IBS, IBS Regional and Non-IBS Events and Meetings

[View the meetings calendar here!](#)

Is something missing? Would you like to add your meeting or event to our calendar? If so, please send an email to [IBS@biometricsociety.org](mailto:IBS@biometricsociety.org) with the following information:

1. Event Title
2. Event Description & Location
3. Event Category (IBS Regional Event, IBS Event, Non-IBS Event)
4. Event Link
5. Start/End Date

## News from the Committee on Communications

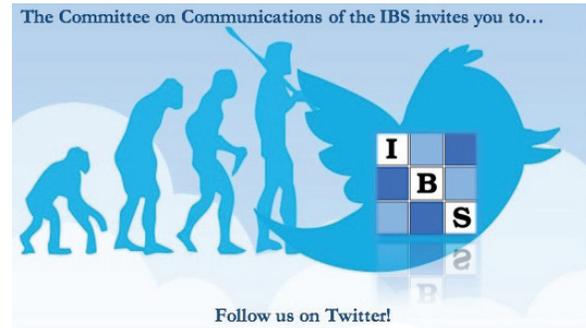
In 2020, the Committee on Communications underwent changes in its members and chair. Kyungman Kim (ENAR) passed the role as chair to Stephanie Roll (German region). In addition, several new members were welcomed. In its first meeting in April, members of the committee had the chance to meet each other as well as discuss upcoming topics. In particular, the new possibilities offered by the IBS Community website were discussed (<https://members.biometricsociety.org/community/community-home>). We were hoping to build this community board with information and useful items for IBS members. We encourage all members to shape and interact within this new platform. In addition, we discussed tagging the IBS on any of the regional posts or news and events to be shared by the IBS social platforms.

Follow IBS on Social Media!

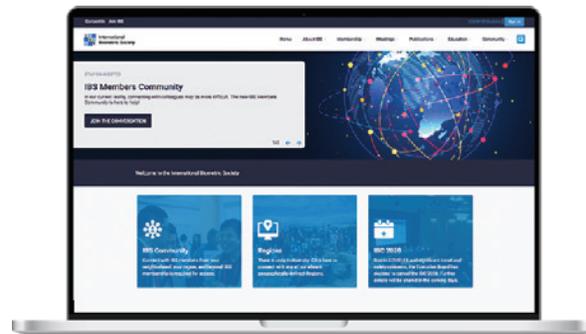
**Twitter** [@IBSstats](#)

**Facebook** [@BiometricSociety](#)

**LinkedIn** [@International Biometric Society](#)



## Get Connected!



Join the IBS Members Community::

1. [Sign in](#) to [biometricsociety.org](http://biometricsociety.org) with your existing credentials. [Contact us](#) if you need assistance.
2. Visit your **profile** to update your personal information and edit privacy settings.
3. Click **Community** in the main navigation and select **All Communities**. Members will see a brand new **IBS Members** community to join.
4. After you join the community, be sure to click on the group **Settings** and customize **Email Notifications** so you can only receive the information you want (real-time updates, a daily digest, or no emails). You may set your preferred **email** address here as well.
5. The final step, **post an update or topic!** This channel is for IBS members, the goal is to empower 6,000 colleagues worldwide to share their unique perspectives, reach out, and connect. It will take some time to cultivate these discussions organically so we appreciate your patience and also your input.

We welcome all comments and feedback about the new website. [Let us know](#) what you think!