Volume 26, No. 2 • Summer 2019

BIOPHARMACEUTICAL REPORT

Chair: Heather Thomas and Richard Zink **Editors:** Ilya Lipkovich and Julia Jingyu Luan

CONTENTS

Note from the editors

Welcome to the second issue of the Biopharmaceutical (BIOP) Report for 2019! This issue's featured article by **Haoda Fu** and **Vipin Gopal** from Advance Analytics and Data Sciences at Eli Lilly and Company presents a succinct and insightful view on data driven personalized solutions.

A question from **Kun Chen** (Gilead) whether To Be or Not To Be a member of the ASA Biopharmaceutical Section has been answered by several ASA BIOP members including graduate students and seasoned industry statisticians.

A note by **Mathew Rotelli** (Eli Lilly and Company) introduces an important topic of Bioethics. **Michael O'Kelly** (IQVIA) and **Russell Reeve** (IQVIA) tell us a tale of statistical consulting in Pharma, as seen from different sides of Atlantic.

In anticipation of the upcoming 40th anniversary of the ASA Biopharmaceutical Section (in 2020) we continue a series of vignettes from some of the key contributers to the section who reflect on the past and offer insights for the future. This issue contains vignettes by **Sue-Jane Wang** (FDA) and **José Pinheiro** (Janssen Pharmaceuticals).

This issue presents updates on Biopharmaceutical Section Working Groups: Safety (**Judy Li, Amit Bhattacharyya** and **Bill Wang**) and Real World Evidence (**Weili He** and **Martin Ho**).

We also provide brief information on the recent 6th Nonclinical Biostatistics Conference and on the upcoming ASA Biopharmaceutical Section Regulatory-Industry Workshop.

We hope you enjoy reading this issue and welcome feedback, suggestions for improvement and topics of interest that you would like to see in the future issues.

IMPROVING PATIENT OUTCOMES THROUGH DATA DRIVEN PERSONALIZED SOLUTIONS

Haoda Fu, Ph.D. and Vipin Gopal, Ph.D., Eli Lilly and Company

Introduction

The idea of improving patient's outcome through data and analytics has been thousands of years old, and it remains a key topic in today's healthcare system. The recorded history of evidence-based healthcare decision goes back to the biblical descriptions in the "Book of Daniel" in The Bible in 500 BC. King Nebuchadnezzar, a resourceful military leader, during his rule in Babylon, ordered his people to eat only meat and drink only wine, a diet he believed would keep them in sound physical condition. But several young men of royal blood, who preferred to eat vegetables, objected. The king allowed these rebels to follow a diet of legumes and water for 10 days. When Nebuchadnezzar's experiment ended, the vegetarians appeared better nourished than the meat-eaters, so the king permitted the legume lovers to continue their diet. This probably was the one of the first times in evolution of human species that an open uncontrolled human experiment guided a decision about public health [1]. In the 18th century, James Lind is considered the first physician to have conducted a controlled clinical trial of the modern era. Dr Lind (1716-94), whilst working as a surgeon on a ship, was appalled by the high mortality of scurvy amongst the sailors. He planned a comparative trial of the promising treatments for scurvy. His trial covers the essential elements of a controlled trial and concluded that oranges and lemons can be used to treat the symptoms [2]. The idea of randomization was introduced in 1923, and the first randomized control clinical trial was conducted by Sir Austin Bradford Hill in 1946 which demonstrated that the efficacy of streptomycin for treating tuberculosis [3]. The greatest influence of this trial lays in its methods which have affected virtually every area of clinical medicine, and, since then, randomized control trials soon became a gold standard for medical research [4]. However, studies on patient's heterogenous response to a treatment has been largely missed, although its importance has been noticed in very early stage. The need of evidence based personalized medicine has been appeared in an early critique of statistical methods in medicine published in 1835 [5]. However, it was not until recently, with the advancement of technology to collect more granular level individual patient information, including DNA sequencing, both clinicians and statisticians had realized a great need to develop statistical methods to advance the concept of a personalized solution [6].

In this paper, we provide a review of the past and present state of personalized solutions in healthcare, along with a discussion of the challenges for future development.

Subgroup analysis, subgroup identification, and modern personalized medicine

Subgroup analysis, subgroup identification, and personalized medicine are closely related, and they can be considered as three generations of methods for personalization, although their objectives could be somewhat different.

Subgroup analysis is often referred to as evaluation of treatment effects on a predefined subgroup of patients based on their baseline covariate values. These subgroups are often prespecified, and such analyses are often specified in the protocol for secondary or exploratory objectives. The purposes for such analysis vary. The sponsor may use subgroup analysis as a salvage strategy for a phase III trial in case it may not meet the primary objective for all enrolled patients. It can be used to pursue an additional treatment indication for a special patient population within a large study. It can also be used to evaluate scientific hypotheses for further studies. Thus, subgroup analysis is utilized for both confirmatory and exploratory purposes. One of the fundamental issues for subgroup analysis comes from multiplicity, which impacts both hypothesis testing and treatment effect estimation. The study often involves various objectives including multiple potential

subgroup analysis. Therefore, to confirm the finding, the p-value has to be adjusted. One commonly used strategy is to adopt a graphical testing method [7]. In addition, the U.S. Food and Drug Administration and European Medicines Agency have released guidance documents that discuss regulatory and clinical and statistical approaches to subgroup analysis [8,9]. Different subgroups can be overlapping, so their test statistics' correlation can be quantified under a null hypothesis. Therefore, further efficiency could be achieved by incorporating their correlation structures. Multiplicity also results in a selection bias in estimating the treatment effect. In exploratory subgroup analysis, we often choose the most significant subgroup for further evaluation. An immediate question is what are the treatment effects. Naively using the observed treatment effects from this selected subgroup will overestimate the treatment effect. A debiased estimation is needed for which empirical Bayesian methods could be used [10].

Explicitly specifying the subgroups before conducting analysis is often challenging. The concept of identifying subgroups from data is attractive. Retrospective data driven subgroup identification has gained significant popularity for the past decade, and various methods have been proposed to search subgroups for hypothesis generation. To highlight a few, Su et al. [11] proposed the interaction trees method which extend classification and regression trees (CART) by incorporating a treatment by split interaction. Lipkovich et al. [12] developed algorithms extending bump hunting methods to search for differential treatment effects. Loh et al. [13] extended their previous work to search for subgroups while adjusting for covariate selection bias when we have both categorical and continuous covariates. Lipkovich et al. [14] provides a comprehensive review of this topic. The notorious challenge of multiplicity still exists and may even become more severe. The total search space is often less understood analytically, which poses additional challenges for adjusting p-values. Some ad-hoc approaches are often adopted, such as splitting data into training and testing datasets (or out of bag samples) to evaluate the estimated subgroups.

Besides multiplicity, there are other fundamental challenges for subgroup identification. First, a unique definition of subgroups does not exist. For example, some methods are intended to maximize the treatment by covariate interaction, and others search for differential treatment effects. Second, many of the existing methods are tree-based approaches, and their optimization is done layer by layer. The final solution may not be the global optimal solution, and their theoretical properties are difficult to evaluate. Third, these methods only focus on treatment benefit. As a consequence, those methods often face a dilemma of whether to select a small subgroup with significant treatment benefit versus a larger subgroup having moderate treatment advantage. Furthermore, the geometric shapes of subgroups are often not clearly defined. Some methods only search for a single rectangle shape subgroup, and some methods allow multiple half open spaces.

The ultimate purpose of subgroup identification is to maximize patient benefit because we believe that patients in such subgroups can achieve better outcomes when taking treatment. By viewing subgroup identification as an outcome optimization, we form a new framework under the individualized treatment recommendation (ITR). ITR is a modern method for personalized medicine and has gained tremendous popularity recently. These methods search treatment assignment rules in a defined functional space (e.g. linear models or tree models) to maximize patient benefit [15]. The method and framework also have significant benefit over the traditional methods in that they can handle both randomized control trials as well as observational studies by adjusting for confounders through inverse probability weighting scheme or doubly robust methods.

Following the work by [15, 16, 17], Fu Zhou and Faries [18] connected subgroup identification problems in pharmaceutical setting with personalized solution methods in academic studies. They reinterpreted the acronym ITR from traditional Individualized Treatment Rule to Individualized Treatment Recommendation to broaden its use and increase acceptance among clinicians. Their paper also proved that for all subgroup identification related methods, it is important to remove the intercept and covariate effects to increase numerical performance, which is similar to having a centralized covariate matrix before fitting a linear model. Their method uses a comprehensive search scheme to maximize a single objective function within a 3-layer tree structure. The authors argued that this setting satisfied majority of the clinical need. An R and C++ implementation of this method can be found at (*https://github.com/fuhaoda/ITR*).

Zhao et al. [16] makes a connection that optimizing the patient outcome can be formulated as a weighted classification problem. This insightful connection links the field of machine learning with personalized medicine and opens up many possibilities. For example, [16] modified support vector machine for ITR, and [19,20] generalized the idea into multicategory treatments with geometric interpretation. Zhang et al. [19] is also the first paper to prove a method with Fisher consistency in selecting an optimal treatment among multicategory choices. Liang et al. [21] modifies deep learning methods by a weighted softmax loss function, so that we can leverage the deep learning framework for more complicated personalized medicine problems when data set are large. Doubleday et al. [22] extends random forest methods for personalized solutions and proposed corresponding variable importance in the ITR setting.

In practice, the users of personalized solutions may not only care about maximizing treatment efficacy but also about drug safety. Therefore, in many situations, we have to consider both. However, the treatment recommendation is a ranking problem which can only be done in the one-dimensional case (directly). In general, we have 3 ways to handle multiple responses. The first approach is the clinical utility index approach so that we can maximize a weighted outcome. The second approach is a constraint optimization approach while we can control safety while maximizing patient benefit [23]. The third approach is to estimate an efficacy-safety trade off through data. In the next sections we will discuss further opportunities and challenges of personalized solutions.

Future research in personalized medicine

The ITR reframes the traditional subgroup identification problems and opens new opportunities for personalized medicine. It connects with machine learning through a weighted classification problem. This approach has also been studied independently in computer science where it was referred to as the contextual based bandit problem [24]. The solution belongs to single step off policy reinforcement learning [25]. It is worth noting that the reinforcement learning algorithms are key algorithms in the field of artificial intelligence. Google DeepMind used reinforcement learning algorithm to develop Alpha Go and Alpha Zero to beat the best human Go game player in 2016. However, application of these algorithms in the medical domain is not straightforward, as there are many challenging issues to address. One question is how we can continue to improve the recommendation engine. Once the ITR algorithm is obtained from a training dataset, it will be a deterministic function conditional on a patient's covariate information. To continue to improve the algorithms, some randomness for treatment exploration has to be introduced. The epsilon-greedy algorithm, Thompson sampling, and upper confidence bounds are three popular choices. However, in medicine, it would not be ethical to allow patients to try solutions which are known to be risky. Therefore, research on how to balance and quantify individual risk, and then building it into the exploration phase are needed.

The reinforcement learning framework greatly extends personalized medicine from a single decision point to multi-stage personalized interventions. For chronic disease, patients often have to switch or intensify their treatments. Dynamic treatment regimes [26] provide statistical interpretation of reinforcement learning. The Q -learning methods based on Bellman equations are popular approaches. Sequential multiple assignment randomized (SMART) and Micro-randomization trials provide a way to formally study and develop algorithms for a personalized solution [27]. Recently [28] extended the traditional dynamic treatment regime from a few stagewise decisions into an almost continuous horizon for mobile health.

Besides adopting reinforcement learning approaches into personalized medicine areas, in medicine there are other unique challenges that needs to be address. For example, consider constraints on diagnostic costs. Suppose we only have $100 to diagnose a sub phenotype for better personalized intervention. We can either choose 10 low cost biomarkers or two expensive lab tests. Under such a constraint, how can we maximize our diagnostic accuracy? The cost can be generalized to convenient cost; in mobile health, it may be unrealistic to ask patients to wear 10+ sensors for personalized interventions. How can we select the most relevant devices based on different patient profile to achieve adequate diagnostic accuracy?

With advancements in devices and technology, we are now able to collect more data to better quantify individual patients. This data provides a great opportunity to generate actionable insights to improve patient outcomes. From the first documented controlled trial in The Bible to the

gold standard randomized control trial is a journey, and this journey continues in the 21th century as personalized medicine. It starts as subgroup analysis to artificial intelligence-based reinforcement learning ITR algorithms. We are confident that this data-driven medical decision-making will continue a center topic in medical research, and it will continue to be an essential way to improve patient outcomes!

Acknowledgements

The authors thank Ilya Lipkovich for his constructive suggestions and edits for this manuscript.

References

1. Collier R. Legumes, lemons and streptomycin: A short history of the clinical trial. CMAJ. 2009;180:23–24.

2. Bull JP. MD Thesis: University of Cambridge; 1951. A study of the history and principles of clinical therapeutic trials.

3. MRC Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. BMJ. 1948;2:769–83.

4. Yoshioka A. The Randomized Controlled Trial of Streptomycin in The Oxford Textbook of Clinical Research Ethics. In: Emanuel EJ, Grady C, Crouch RA, Lie RK, Miller FG, Wendler D, editors. Oxford: University Press Oxford; 2008. pp. 46–60.

5. Poisson, Messrs. "Statistical research on conditions caused by calculi by Doctor Civiale." International Journal of Epidemiology 30, no. 6 (2001): 1246-1249.

6. Longford, Nicholas T., and John A. Nelder. "Statistics versus statistical science in the regulatory process." Statistics in medicine 18, no. 17-18 (1999): 2311-2320.

7. Bretz, Frank, Willi Maurer, Werner Brannath, and Martin Posch. "A graphical approach to sequentially rejective multiple test procedures." Statistics in medicine 28, no. 4 (2009): 586-604.

8. Food and Drug Administration. Guidance for industry: enrichment strategies for clinical trials to support approval of human drugs and biological products, 2012.

9. European Medicines Agency. Guideline on the investigation of subgroups in confirmatory clinical trials, 2014.

10. Efron, Bradley. "Tweedie's formula and selection bias." Journal of the American Statistical Association 106, no. 496 (2011): 1602-1614.

11. Su, Xiaogang, Chih-Ling Tsai, Hansheng Wang, David M. Nickerson, and Bogong Li. "Subgroup analysis via recursive partitioning." Journal of Machine Learning Research 10, no. Feb (2009): 141-158.

12. Lipkovich, Ilya, Alex Dmitrienko, Jonathan Denne, and Gregory Enas. "Subgroup identification based on differential effect search—a recursive partitioning method for establishing response to treatment in patient subpopulations." Statistics in medicine 30, no. 21 (2011): 2601-2621.

13. Loh, Wei-Yin, Xu He, and Michael Man. "A regression tree approach to identifying subgroups with differential treatment effects." Statistics in medicine 34, no. 11 (2015): 1818-1833.

14. Lipkovich, Ilya, Alex Dmitrienko, and Ralph B D'Agostino Sr. "Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials." Statistics in medicine 36, no. 1 (2017): 136-196.

15. Qian, Min, and Susan A. Murphy. "Performance guarantees for individualized treatment rules." Annals of statistics 39, no. 2 (2011): 1180.

16. Zhao, Yingqi, Donglin Zeng, A. John Rush, and Michael R. Kosorok. "Estimating individualized treatment rules using outcome weighted learning." Journal of the American Statistical Association 107, no. 499 (2012): 1106-1118.

17. Zhang, Baqun, Anastasios A. Tsiatis, Marie Davidian, Min Zhang, and Eric Laber. "Estimating optimal treatment regimes from a classification perspective." Stat 1, no. 1 (2012): 103-114.

18. Fu, Haoda, Jin Zhou, and Douglas E. Faries. "Estimating optimal treatment regimes via subgroup identification in randomized control trials and observational studies." Statistics in medicine 35, no. 19 (2016): 3285-3302.

19. Cong Zhang, Jinxiang Chen, Haoda Fu, Xuanyao He, Yingqi Zhao, and Yufeng Liu (2019),``Multicategory Outcome Weighted Margin-based Learning for Estimating Individualized Treatment Rules," Statistica Sinica, 10.5705/ss.202017.0527.

20. Qi, Zhengling, Dacheng Liu, Haoda Fu, and Yufeng Liu. "Multi-Armed Angle-Based Direct Learning for Estimating Optimal Individualized Treatment Rules With Various Outcomes." *Journal of the American Statistical Association* (2019): 1-33.

21. Muxuan Liang, Ting Ye, Haoda Fu (2018), ``Estimating individualized optimal combination therapies through outcome weighted deep learning algorithms," Statistics in Medicine. 2018;1--18. *https:// doi.org/10.1002/sim.7902*

22. Kevin Doubleday, Hua Zhou, Haoda Fu, Jin Zhou (2018), ``Generation of Individualized Treatment Decision Tree Algorithm with Application to Randomized Control Trials and Electronic Medical Record Data", *Journal of Computational and Graphical Statistics,* doi:10.1080/10618600.2018.1451337

23. Yuanjia Wang, Haoda Fu, Donglin Zeng (2018), ``Learning Optimal Personalized Treatment Rules in Consideration of Benefit and Risk: with Application to Treating Type 2 Diabetes Patients with Insulin Therapies," *Journal of the American Statistical Association,* Volume 113, 2018 - Issue 521 DOI:10.1080/01621459.2017.1303386.

24. Li, Lihong, Wei Chu, John Langford, and Robert E. Schapire. "A contextual-bandit approach to personalized news article recommendation." In Proceedings of the 19th international conference on World wide web, pp. 661-670. ACM, 2010.

25. Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT press, 2018.

26. Murphy, Susan A. "Optimal dynamic treatment regimes." *Journal of the Royal Statistical Society:* Series B (Statistical Methodology) 65, no. 2 (2003): 331-355.

27. Klasnja, Predrag, Eric B. Hekler, Saul Shiffman, Audrey Boruvka, Daniel Almirall, Ambuj Tewari, and Susan A. Murphy. "Microrandomized trials: An experimental design for developing just-in-time adaptive interventions." Health Psychology 34, no. S (2015): 1220.

28. Luckett, Daniel J., Eric B. Laber, Anna R. Kahkoska, David M. Maahs, Elizabeth Mayer-Davis, and Michael R. Kosorok. "Estimating dynamic treatment regimes in mobile health using V-learning." *Journal of the American Statistical Association* (2019): 1-34. .■

WHY BIOPHARMACEUTICAL SECTION OF ASA?

Kun Chen (Gilead)

Are you a member of the American Statistical Association (ASA)? Would you like to learn more about applied statistics in developing drugs, biologics and devices from both industry and regulatory aspects? As the Chair of Membership Committee of the Biopharmaceutical (BIOP) Section, I would like to invite you to join us. Find out why you should join us from the following general description of BIOP section and the personal experience shared by both executive and junior members.

The BIOP section overview

BIOP Section was initially a subsection of the Biometrics Section starting in 1967, and became a full section in 1981. BIOP focuses on the application of statistics to the development and use of therapeutic drugs, biologics and devices in humans and animals. The BIOP Executive Committee consists of members that help drive many subcommittees and activities including Membership, FDA-Industry Workshop, Student Paper Award, Contributed Paper and Poster Awards, Distance Learning, and Fellows Nominations, as well as supporting a variety of activities such as Scientific Research Working Groups, Web-based Training Series, Statistical Outreach Program, Student Paper Competition, Best Contributed Paper and Poster Competition, annual Joint Statistical Meeting (JSM) Mixer, Regulatory-Industry Statistics Workshop (RISW), biennial Non-Clinical Biostatistics Conference (NCB), bi-annual BIOP report, and a Membership Survey.

BIOP section offers training including webinars, YouTube Channel videos, full- and half-day online courses, and short courses at JSM, RISW and NCB. The annual RISW helps people gain industry and regulatory perspectives on the challenges of development in drugs/biologicals/devices and gain access to very affordable training. The Biopharmaceutical Report publishes articles on key biostatistical topics, general updates and summaries of Section meetings, and information on conferences sponsored by BIOP section. Podcasts and YouTube videos highlight statistical leadership and specific statistical topics important to our industry. Scientific working groups conduct research on current challenges and advanced statistical methodologies.

SOME PERSONAL EXPERIENCES

ERIK PULKSTENIS, *Council of Sections Rep (2017-2019), VP Data and Statistical Sciences, Abbvie Inc*

The Biopharmaceutical Section is the first place to go to network in the field. Through interactions with members and many leadership opportunities, the Section is an amazing place to connect with new and old colleagues, and engage in industry/regulatory/academic conversations. The thing that I love about the Section is that it is a great way to stay aware of emerging trends in our discipline. The community is vibrant and engaged and if one wants to grow their network or mentor/be mentored, there is no better way to do so.

WILL EAGAN, *PhD Candidate in the Department of Statistics at Purdue University*

I am developing novel methodology for antimicrobial susceptibility testing for my Ph.D. dissertation. After graduation I desire to work in the pharmaceutical industry. I have received tremendous encouragement from the Biopharmaceutical Section. Seeking both funding to attend JSM and recognition, I applied for the 2018 Biopharmaceutical Section Scholarship Award. To my delight I was named one of the three winners. The award money enabled my travel to the 2018 JSM in Vancouver. There I attended the student leadership center to represent and to lead Purdue's eventual co-winning team in the student leadership challenge, attend presentations relevant to my research, and talk with those experienced in my dissertation topic. When I attended the open business meeting for the Biopharmaceutical Section, I was blown away by the amount of career insights offered by the section members. Upon returning home, I immediately joined the Biopharmaceutical section. I look forward to attending the section's open business meeting at every future JSM.

YONGMING QU, *BIOP Publication Officer-Elect (2019-2021) Senior Research Advisor, Eli Lilly*

I attended JSM in 2001 for the first time when I was still a graduate student from Iowa State University. I felt that

I was totally lost in the conference. In 2004, I attended JSM again at Toronto after I joined Eli Lilly. I found the topics relevant to my work were mostly from sessions sponsored by Biopharmaceutical Section. When travelling back home, at the airport I met a statistician from another company who was a Biopharmaceutical Section member. Through the conservation, he had shared with me some insights and gave me mentoring advice that contributed tremendously to my productivity and benefited my career. Since then I became an active Biopharmaceutical Section member and volunteered for organizing the Poster Award, publishing the Biopharmaceutical Report, and recently as the Publication Officer. I really like the opportunities that Biopharmaceutical Section offers: a great network for statisticians in the pharmaceutical industry, JSM sessions organized Biopharmaceutical Section, the high-quality webinars (free for members), the very popular Regulatory-Industry Statistics Workshop. Looking back, I wish I had joined the Biopharmaceutical Section earlier!

THEVAA CHANDERENG, *PhD Student University of Wisconsin*

I feel honored to be part of BIOP member. My very first graduate school award was awarded by BIOP and I am extremely thankful to the section for that. The diversity in the section, which includes people from academia and industry, thus providing views from two different perspectives, have always motivated me to be part of BIOP. The working groups dealing with different challenges in drug development have also captured my interest. Besides that, BIOP also provides a great platform to discuss and communicate various biopharmaceutical related problems. Recently, the podcasts have grabbed my attention. The podcasts do not only discuss new interventions in the biopharmaceutical world but it also provides in-depth detail on the methodology and practical application. I would highly recommend students and young researchers to join the BIOP not only to enhance their networking skills but also to "kick start" their career in the area of biopharmaceutics.

KUN CHEN, *Chair of Membership Committee, Sr. Director, Biostatistics, Gilead Science Inc*

I have been working as a statistician at pharmaceutical industry for almost 18 years and as an ASA and BIOP member for many years. As a BIOP member, I have advantages to communicate with my fellow statisticians regarding statistics questions and challenges, find solutions with the help from other statisticians, share the information on upcoming webinars, online trainings, and conferences. With the BIOP membership, I get a registration discount for online trainings, RISW and other meetings. RISW is my favorite statistics conference where I organized and chaired several sessions, networked with statisticians from both industry and regulatory colleagues, learned the implementation of the advanced methodologies in clinical development such as adaptive design, enrichment design, etc.

If you are a member, thank you for joining us and further enjoy the full benefits of this membership. If you not a member yet, we would love to have you aboard so you could take advantage of the BIOP membership to explore future opportunities in your career. If you know someone who is not a member, encourage them to sign up. Membership is $7 for professionals and free for students. ■

BIOETHICS AND STATISTICS

Matthew D. Rotelli, Ph.D., Eli Lilly and Company

What is bioethics, and why is it relevant to statisticians? More importantly, why are statisticians important for bioethics? Let's start with ethics. Ethics is the discipline dealing with moral duties and obligations, and, more generally, what is good or bad about different actions one might take. Ethics are influenced by culture, societal norms, and belief systems. It is natural to find variability in what is considered acceptable, admirable, or distasteful between individuals or between any groupings of individuals. Still, certain principles emerge that enable us to assess our behavior within our communities. For example, the ASA has developed Ethical Guidelines for Statistical Practice (*https://www.amstat.org/ASA/About/Ethical-Guidelines-for-Statistical-Practice.aspx*) which are required to be upheld for PStat® and GStat accreditation and are expected to be upheld by every member of ASA, any other practitioner of statistics, and their employers.

Bioethics is the consideration of ethical issues in biological research and applications, such as medicine. You can see where this would be particularly relevant for the pharmaceutical industry, where animal and human research studies must be conducted to discover, develop, and market medicines. Bioethics as a discipline is relatively new, emerging as a result of highly publicized cases of abuses and atrocities committed by researchers. Foundational papers include the Nuremberg Code (from the Nazi Doctors' Trial), the Declaration of Helsinki (from the World Medical Association), and the Belmont Report (in response, in part, to the Tuskegee Syphilis Study). More recently, the Council for International Organizations of Medical Sciences (CIOMS) has updated their guidelines entitled "International ethical guidelines for health-related research involving humans." Similarly, the United States Health and Human Services Department Office of Human Research Protections updated the Common Rule governing all federally funded research. This is a reflection of the ever-evolving nature of the field of bioethics in response to advances in technology including genetic testing and digital health applications, for example, as well as increasing concerns related to privacy protections.

Some of the principles that emerge from these documents include respect for persons, beneficence, nonmaleficence, and justice. Respect for persons requires ensuring that research participants have the right to fully exercise their autonomy: ensuring that those who are capable can provide voluntary informed consent, and ensuring that protections are in place for those who are not capable. Beneficence requires that researchers seek to maximize benefits and minimize harms. Nonmaleficence requires that researchers inflict no harm. Justice requires that research procedures be carried out fairly and equitably and that there is a fair distribution of costs and benefits to all potential research participants. Through the application of these principles, bioethics plays a role in study design, choice of comparators, informed consent process and content, selection of countries and sites, requests for access to investigational treatments outside of clinical trials, animal care and use, handling of special populations (e.g. pediatrics), and timing and content of research publications, just to name a few common areas.

Based on these principles, Eli Lilly and Company developed and published "Eli Lilly and Company's bioethics framework for human biomedical research."[1] This framework contains 13 essential elements for ethical human biomedical research, some of which readily extend to animal research, re-use of data (secondary research), and handling and re-use of human biological samples. I am privileged to lead the Lilly Bioethics Program, where our dedicated staff and advisors consult with teams to help them navigate decisions when challenging and often conflicting bioethical considerations arise. I am proud to work for a company that prioritizes bioethics and has developed positions on topics ranging from stem cell research to choice of control for clinical trials to scientific publication (*https://www.lilly.com/bioethics*). This commitment helps the organization keep focus on patient well-being and protecting our research participants.

Two fundamental elements of a bioethics framework for research are scientific validity and social value. If a research study or project is not scientifically valid,

it cannot be justified ethically, and the other elements become irrelevant. Likewise, research must have social value – the results will contribute to generalizable knowledge and any products developed will directly or indirectly benefit public health and well-being.

Statisticians have deep expertise in assessing the scientific validity and generalizability of research. They are capable of incorporating design and analytic approaches that maximize these elements. They are trained in determining that a study may fall short and therefore should not be conducted. They are knowledgeable in summarizing key aspects of design, methodology, and limitations so that others can assess the strength and applicability of the research results. Thus, statisticians are essential to the conduct of ethical research.

So, I hope I have sparked your interest in bioethics. Statisticians are critically important for the effective incorporation of these principles into research, from design to publication. All scientists, including statisticians, are responsible to ensure that the research they conduct is ethical. Whenever something doesn't feel right, speak up! The resulting conversations will either clarify the ethical rationale for the current research proposal or change the direction of the research. Either way, the research team will have more confidence in the justification for their work. Better research will lead to increased public trust in the research results. This trust will result in both increased participation in research as well as better adherence and outcomes for patients. That is ultimately why we do the work we do. ■

1. Luann E. Van Campen, Donald G. Therasse, Mitchell Klopfenstein & Robert J. Levine (2015) Eli Lilly and Company's bioethics framework for human biomedical research, Current Medical Research and Opinion, 31:11, 2081-2093, DOI: 10.1185/03007995.2015.1087987

THE CHALLENGES AND REWARDS OF STATISTICAL CONSULTING IN PHARMA: TWO VIEWS ACROSS ATLANTIC FROM THE SAME COMPANY

Michael O'Kelly, Ph.D., IQVIA and Russell Reeve, Ph.D., IQVIA

Michael O'Kelly, (Ireland)

For me, statistical consulting has consisted of a) solving statistical problems for company colleagues; b) solving statistical problems for our clients; and c) research into statistical methodologies that could improve what we do in the drug development industry (often inspired by a) and b)). Items a) and b) include the training in statistical methodologies that I provide for company colleagues, including non-statisticians, and for company clients.

"Solving statistical problems" covers a multitude: much of the most useful problem-solving starts with a small but urgent query from a colleague. Often, in these cases, I can find a solution that was already published by somebody else. Sometimes, I will know the solution because the problem has come up before; sometimes, I will find it one of the treasured books in the library cupboard down the hall; and often Google will help. An example came in the other day: "In estimating the subgroup efficacy score for a treatment group, what values should I give the baseline prognostic variables?". On that occasion, examples of approaches to take came from some code shared long ago by Professor James Roger, and from a paragraph by ex-colleague Dr Sonia Davis in a book on which we had collaborated.

On rare lucky days I may even think of a new idea to solve the problem, or maybe put two old ideas together. A recent example of this was a request to find a measure of representativeness when sampling for a survey of drug use. Working with my colleagues doing the surveys, it seemed that one could use a proposed sample as starting values and provide an effective set of improvements to the representativeness of the sample using the ordinary stratified chi-square statistic to assess the similarity of the distribution of the proportions in the sample vs. the in the target population – not very advanced statistics, but a nice use of the old Cochran-Mantel-Haenszel test. (After I drafted this I found a whole stream of research in this area that was new to me, and which went back to the beginning of the last century.)

Sometimes, a colleague or client simply does not like the idea or solution I present; this is a fine and delightful challenge for the consultant. If the colleague or client and I can come to an agreement despite this mismatch of ideas then the challenge becomes a reward; but sometimes a challenge remains a challenge, and that's a challenge! I won't provide examples of these challenges, but it is important for us statisticians as professionals to understand that challenges that remain challenges (often known outside the business world as "failures") are one's own responsibility. It's not pleasant: one can learn from failures, as the cliché goes, but only because they are indeed failures. They are part of the picture, part of science, part of our lives as consultants.

On the other side of the coin, a solution to a problem will often evolve as I and other statisticians talk; that is probably the best reward a statistical consultant gets. A recent example of this was a half-idea I had for nonparametric imputation of times to event, which evolved into an elegant solution after discussion with the editor of this journal, Ilya Lipkovich (see our eventual publication with Bohdana Ratitch in Pharmaceutical Statistics, Lipkovich et al., 2016).

A final challenge for the technical statistical consulting function in a company is the challenge of accounting financially for the value of the function. The kind of statistical consulting I have been describing often by its nature consists of short bursts of work; and it can stretch the accounting bureaucracy to ensure that the monetary value of the work, and therefore the value of the consulting function, is recognized in money terms.

In summary, the three biggest challenges for statistical consulting in my experience are 1) documenting one's contributions in a form that can be used to justify the consulting in money terms; 2) reconciling or modifying the possible statistical solutions so as to

satisfy actual requirements of the project and the human preferences of the people with the problem; and the most pleasant challenge, 3), is actually using one's statistical brain to find the solutions. The three best rewards are 1) Self: "Wow, I've just thought of a great idea to solve the problem"; 2) Colleague: "OK, but have you thought of trying this?"; 3) Client: "Thanks, that's perfect!".

Russell Reeve, (USA)

It is interesting to reflect on what we do, versus what other people think we do. Right after I graduated with my PhD, a good friend asked why it would be interesting to work long hours on a computer tabulating numbers. Of course, that is not what statisticians do, in fact they generally work with other people rather than with computers. Indeed, as a consultant I find myself in meetings or working with other people one-on-one a majority of every day. And this is where it gets interesting, as I have found the most rewarding aspect of the job is solving difficult, important problems with people of different skillsets and knowledge.

Let's take an example from the 1990's. Mycophenolate mofetil (MMF) is a compound that was being developed to reduce rejection episodes in renal transplant patients. At the time, rejection episodes were a significant risk, and the therapeutic agents reduced rejection episodes in only a fraction of the patients, and hence a new therapeutic agent would be useful. But the chief challenge was that if a rejection episode occurred in a patient, then the circulating concentration of the MMF would be high. This was because MMF was cleared through the kidney. If one experienced a rejection episode, then the kidney would not perform properly and hence the drug was not excreted. When comparing patients with rejection episodes to those without, the MMF concentration would be higher, which could lead to the conclusion that higher concentrations of MMF yielded worse efficacy outcomes. Hence, a new approach was needed to investigate the effects of concentration on efficacy outcomes, from a causative instead of correlative perspective. The idea? Utilize a new study concept that Laszlo Endrenyi had developed (Controlled Clin Trials 1991;12:780-94), where the patients are randomized to drug concentrations instead of doses, and the doses are adjusted to get to the appropriate concentrations. Mike Hale, my supervisor at the time, and I developed the trial design by (1) working with the medical team to ensure that the concept would work medically, (2) obtained approval from regulatory to discuss with the FDA regulators, (3) developed the exposure-efficacy models with the pharmacokinetics group, (4) developed a trial simulation, and (5) estimated the number of tablets of various sizes that needed to be manufactured of each size. The simulations crashed the corporate computer system, which meant we had to work methods of reducing the computer load. We see from this some aspects of statistical consulting in the pharmaceutical industry:

- Need to work with subject-matter experts as a team
- Develop habit of learning the subject-matter
- Be willing to speak up, and need to be willing to suggest solutions
- A project may require working with many different colleagues
- The statistical aspects may be only a small part of the project, and we need to ensure that the statistical methods support the project
- Unanticipated challenges may appear, some having nothing to do with the project itself
- Need to be flexible and solution-minded

Engaging in these behaviors has been rewarding. I have learned how to run an bioassay, watched manufacturing and packaging, and learned more medicine than I ever expected.

I have long ago gave up on expecting my carefully laid plans for the day to work out. In consulting, the work is often urgent, spontaneously erupting through email, arriving in a M/G/1 queue, and often requires novel solutions. I have found this process of arriving work to be interesting and even useful at time. True, there are times when overtime is required to satisfy project deadlines, especially with multiple projects running in parallel. But it also means some downtime where I can develop new skills, write a paper, or study some new area of medicine that I will need.

So what do I do? I do not analyze data very frequently. Indeed, statistics is not about data analysis at all, but rather is about understanding processes with a stochastic component. What I do is analyze clinical trial processes, describe the distribution of possible outcomes, and propose methods for answering questions.

And one of those questions that invariably come up during a trial design process is: "What is the sample size?" Most non-statisticians think of this as a pretty cut-and-dried question, but is in fact an ill-defined question. Sample size to achieve what? What is the objective of the study? What should the objective be? Is the currently proposed design well-equipped to satisfy the objective, or should we modify it? We might be able to reduce the size of the trial if we introduce a non-responder washout period and then use a randomized withdrawal design in the second phase. Should we power at 80% or 90%, and do we really expect the treatment effect to be the point at which we power? Or maybe we should size the trial to maximize the net present value of the asset instead. Indeed, while many project managers are concerned with the cost of the trial, typically the time value of money affects the value much more significantly than does the direct cost.

All of these questions and more come up in consulting. The real process is working with colleagues to solve real problems. It is working with other people that makes this such an interesting and fulfilling occupation. And finally, it is really satisfying to know that my work can have significant effects on the lives of others, most of whom I will never know. ■

A BRIEF OVERVIEW OF PHARMACOGENOMICS AND MEDICAL IMAGING PRECISION MEDICINE

Sue-Jane Wang, Ph.D.,* US Food and Drug Administration

The advent of medical technology and bioinformatics has been facilitating active research and development that builds its ability to target molecular genomics of individual subjects. Applications to individualization of medical treatment (also known as precision medicine) have been enthusiastically on the lookout. Since late 1990, pharmacogenetics has gradually resurfaced from research efforts in the mid-20th century. In pharmaceutical developments, an exciting example is the realization of pharmacogenomics. In fact, research communities distinguish between pharmacogenetics and pharmacogenomics. At one point, the distinction was made such that studying subjects at DNA level belongs to pharmacogenetics, and pharmacogenomics is a clinical science studying subjects at RNA level or molecular level. In what follows, pharmacogenomics will be used.

To detect the presence of a (genomic) biomarker, a diagnostic assay is required. Pharmacogenomics in drug development, therefore, relies upon the development of an in vitro diagnostics (IVD) to identify individuals with the (genomic) biomarker of potential therapeutic advantage and development of a targeted experimental therapeutics. The molecularly targeted drug development framework has embraced a co-development approach with drug administration accompanying with a specific IVD and a separate development approach without the need of simultaneous approval of therapeutics and clearance of an IVD. Irrespective of the approaches, biomarker has been the mainstay and plays a key role in bridging between the IVD and therapeutics in their developments.

A molecular genomic biomarker, assessed by an IVD and depending on its context of use, can have a variety of clinical utilities. These utilities include pre-therapeutics, e.g., risk susceptibility, prognostic, and diagnostic, and post-therapeutics (monitoring, response, reasonably likely surrogate biomarker endpoint); see *https://www.ncbi.nlm.nih.gov/books/NBK326791/*. Several therapeutics have been approved relying on molecular biomarkers, e.g., ivacaftor is indicated for the treatment of patients with cystic fibrosis (CF) who have a G551D mutation in the CFTR gene (ref: *https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/203188s019lbl.pdf*). Genetic testing is required prior to initiating treatment with ivacaftor if a patient's CFTR genotype is not known. For a list of pharmacogenomic biomarkers seen in drug labeling, see *https://www.fda.gov/drugs/science-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling* (accessed on 6/30/2019).

Statisticians' contributions to the field of pharmacogenomics, molecular biomarker, IVD-therapeutics development have been active and can be summarized in three aspects: (i) statistical methodologies including (novel) study designs, (ii) statistical analysis approaches to assessing efficacy, safety and benefit-risk evaluation, and (iii) the movement of pharmacogenomics incorporating biomarkers for drug development has also motivated the transformation from one-size-fits-all evaluation approach to precision medicine taking into account genomic variation among individuals and aiming at individualization of treatment. In principle, computing overall average treatment effect assumes that there is one and only one statistical distribution for all subjects studied. The treatment effect can be inferred for the studied patient population. This is often referred to as a one-size-fits-all approach. In contrast, the molecular-target based pharmacogenomics drug development brings more design varieties for consideration, e.g., biomarker enrichment design, biomarker-based patient selection, biomarker-based treatment selection.

In recent years in drug development, some radiopharmaceutical imaging drugs began to show their clinical utility as patient selection mimicking a diagnostic biomarker. The recent approval of lutetium Lu-177 dotatate is an example, which used an approved radiopharmaceutical imaging drug, indium In-111 pentetreotide kit, to screen patients via scintigraphy imaging

for identification of patients who have neuroendocrine tumor (NET) bearing somatostatin receptors (SSTR+). Eventually, lutetium Lu-177 dotatate was approved as a radiopharmaceutical treatment of SSTR+ gastroenteropancreatic (GEP) NETs in adults. The lutetium (Lu-177)-dotatate example has shown that one approved radiopharmaceutical drug (Indium In-111 Pentetreotide kit) allowed for targeting of tumors in GEP-NET positive patients and its labeling isotope Lu-177 can deliver therapeutic radiation.

The concepts demonstrated in lutetium Lu-177 dotatate and In-111 pentetreotide kit have laid the ground work for development of radiopharmaceutical combination, known as theranostics, with one product as an imaging drug and the other product as a therapeutic agent. A radiopharmaceutical imaging drug can have other clinical utilities, e.g., monitoring treatment response. When a radiopharmaceutical imaging drug is in its experimental stage, should development of a radiopharmaceutical therapeutic drug wait until the approval of a radiopharmaceutical imaging drug? Can they be pursued in parallel with leveraging in therapeutic trials? More experience is yet to be gained for radiopharmaceutical imaging drug that is in its experimental stage. Depending on the clinical utility of a radiopharmaceutical imaging drug in view of imaging precision medicine, it is foreseeable that more novel study designs, statistical methods and statistical analysis approaches are desirable. Statistical researchers are encouraged to make a timely impact. ■

*Disclaimer: This article reflects the views of the author and should not be construed to represent the views or policies of the U.S. Food and Drug Administration.

LOOKING BACK, MOVING FORWARD

José Pinheiro, Janssen Pharmaceuticals

When I was asked to write a vignette about my experiences as a biostatistician working for many years in the pharmaceutical industry, I was struck with a mix of nostalgia and amusement: it has been a long, highly rewarding path, resembling a random walk with a hint of a purpose. I came to statistics, and later biostatistics, via engineering, which eventually I found out to be the path of many statisticians I came across in my professional life. Engineering provided me with a solid foundation of mathematics that proved to be critical in my later career, but, perhaps more importantly, a passion for solving problems applying quantitative methods that eventually guided me to statistics. This passion still drives me to this very day.

The breadth of application of statistics has always been highly attractive to me, but also a source of anxiety. By pursuing a career in a discipline that can, and should, be used in areas as varied as drug development, telecommunications, and marketing, how does one develop a professional identity the way, say, a medical doctor or a chemist do? For me, the epiphany came during my graduate school days at the University of Wisconsin – Madison, when I was exposed to this definition, by the late Prof. George E.P. Box, of statisticians as the "guardians of science." One would think that science has such a remarkable track record of success in recent human history, that it does not need any nerdy guardians to protect it. However, those of us who have been in the statistical profession long enough have inevitably come across instances when we needed to speak up to calibrate overoptimistic interpretations of results. I believe this healthy skepticism about ad-hoc findings gives us a professional identity, but also makes us, on occasion, not the most popular of project team members in the room. We should embrace it proudly.

Although the path leading me to biostatistics in drug development was not always linear, or even logical, in hindsight, it made sense given my interests and opportunities. I was very fortunate to have strong mentors who guided professionally at critical junctions in my life. I was introduced to clinical biostatistics by Prof. Dave DeMets who taught me, among other things, the importance of what we do as statisticians to protect patients and promote good science. During one of the many late afternoons we spent at O'Hare Airport waiting for our flight back to Madison following a day of meetings with a Data Monitoring Committee, he said that once you are bit by the "biostat bug" it is hard to recover. This turned out to be prophetic in my future career.

Working as a biostatistician in drug development has been a wonderful journey because of the incredible scientists and professionals with whom I have had the privilege to collaborate. I have been part of teams that helped bring transformational treatment to patients, as well as others that endured the bitter disappointment of undelivered promises. Staying true to the data and to science, more broadly, is what makes us statisticians critical members of any project team. As we navigate into a new era of massive amounts of information being continuously collected from previously unimaginable sources, with the challenge of drawing useful knowledge from it, I see it clearly as the dawn of a golden age of statistics. The future is bright for our profession and I am sure that the new generation of incredibly talented statisticians making their way through industry, academia and government will carry on the torch of "guardians of science" with pride and purpose. ■

INTERDISCIPLINARY DRUG SAFETY EVALUATION AND QUANTITATIVE SAFETY MONITORING

AN UPDATE FROM THE ASA BIOP SECTION SAFETY WORKING GROUP

Judy Li[1], Amit Bhattacharyya[2] and William (Bill) Wang[1]

[1]Co-chairs, ASA Biopharmaceutical Section Safety Working Group (ASA-BIOP SWG)

[2]Lead, Outreach & Communication Team, ASA-BIOP SWG

Safety assessment, monitoring and safety surveillance, also referred to as Pharmacovigilance, is a key component of pharmaceutical product development life cycle. Pharmacovigilance is defined as "the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any drug-related problem" [WHO, 2002]. In the past, most pharmacovigilance departments at biopharmaceutical companies focused on the handling of individual adverse event reports (called individual case safety reports [ICSRs]). In recent years, there has been a shift in focus from individual cases towards aggregate analysis to identify potential adverse drug reactions. While individual adverse event case handling remains relevant for understanding specifics of each case, the expectations for marketing authorization holders and clinical trial sponsors have increased in the areas of aggregate safety evaluation in drug development, life cycle management and post-marketing surveillance.

Statistical methodology for safety assessment for monitoring and reporting will need to be further developed to match that for efficacy [O'Neill, 2002, Wang et. al., 2017]. In addition to evolving improved methodologies to support safety evaluation, our new complex world will require a multidisciplinary approach to address this effectively. No longer will the 'safety doc' be alone in determining and explaining safety issues with the simple analysis used in years past. That said, statisticians also need to enhance their knowledge of the clinical aspects and safety background so that they can interpret the safety concerns in drug development. What's needed is a true medical-statistical joint venture in order to develop the models in the 21st century for safety evaluation.

The industry has been slowly preparing itself for such a collaborative approach in the last few years. In 2014, the American Statistical Association (ASA) Biopharmaceutical Section started a safety working group including members from both regulatory agency and industry. The working group initially focused on design and analysis of cardiovascular (CV) safety outcome trials for Type II diabetes drugs, and later expanded into a systematic review of multi-source safety data and corresponding analysis strategies. A few recent publications [Izem et. al., 2018; Ma et. al., 2018; Marchenko et. al., 2017; Marchenko et. al., 2017; Zink et. al., 2017], including a mini-series in the Therapeutic Innovation & Regulatory Science (TIRS) journal in 2018 summarize the work of those initiatives [Seltzer et. al., 2019].

In parallel, another dedicated working group was formed in 2015 to further empower the biostatistics community in the field of quantitative safety monitoring. One initiative of this safety monitoring working group was to focus on a systematic review of statistical methodologies on safety monitoring [Wang et. al., 2017], which include Bayesian and frequentist methods; blinded versus unblinded safety monitoring; individual case analysis versus aggregate meta-analyses; pre-marketing versus post-marketing evaluation; static versus dynamic safety reviews; as well as methods of safety data visualization. Another initiative was to perform a thought-leader interview and industry survey on the current practices and future direction of statistical safety statistics practice, tools, and methods. In addition, a systematic review of safety regulation both at global and regional levels (e.g., US, EU, Japan, China) was also conducted and published [Ball et. al., 2019].

To cultivate interdisciplinary collaboration, the aforementioned two efforts have been integrated and expanded into one joint interdisciplinary working group between the ASA biopharmaceutical section and the DIA scientific communities in 2019. The collaboration between the ASA and DIA offers a great opportunity for cross-functional global innovations. In service of this cause, the team has assembled industry, regulatory and academic experts in the area of drug safety and statistics to develop a series of contributions. These include, but are not limited to

1. Develop interdisciplinary frameworks for aggregate safety assessment planning, and visual tools to enable/enhance cross-disciplinary collaboration (Workstream 1, aka, WS1)

2. Deep dive into various safety assessment and monitoring methodologies, including safety enabled benefit risk evaluation and machine learning methods (WS2)

3. Investigate design/analysis approaches for the integration and bridging randomized controlled trials and real-world evidence for safety decision making (WS3)

The working group has been very active and productive in the last few years in pursuit of strategic and methodological advances in bringing clinical safety assessment, monitoring and reporting as important aspects of patient safety in the clinical trials. Using 2018 as an example, the working group has:

- expanded WS1 into a fully operational and highly productive ASA-DIA working group in Safety evaluation;
- started an interactive safety graphics (ISG) interdisciplinary task force to develop fit-for-purpose visualization tool;
- established the safety-enabling-benefit-risk task force and contributed to the DIA get-the-question-right series;
- established a new WS3 in "Integrating RCT/RWE for safety decision making";
- grew into a sizable multi-disciplinary working group with ~ 40 statisticians and ~10 physicians across industry, regulatory and academia researchers;
- presented at 6+ scientific conferences; delivered 12+ presentations and 3 short courses; 3+ manuscripts were published;
- established leadership council (LC) to enhance our procedure/guideline on how to enroll membership, how to work together, and how to publish together.

The working group had a great start in 2019, with the following new activities:

- A "Benefit Risk Assessment Planning (BRAP)" task force was formed, building on our "Aggregate Safety Assessment Planning (ASAP)" task that started in the earlier year;
- An interactive visualization tool was developed to evaluate drug-induced serious hepatotoxicity (eDISH) by our ISG task force;
- A new safety paper series with the DIA TIRS journal [Seltzer et. al., 2019] has started. In the near term, the paper series will include
  - A summary of our industry survey on safety monitoring [Colopy et. al., 2018];
  - A review of global safety regulatory landscape on aggregate safety assessment [Ball et. al., 2019]; and
  - A framework for aggregate safety assessment planning, currently manuscripts under development.
- A book project "Quantitative Methodologies and Interdisciplinary Practice for Safety Monitoring and Benefit Risk Evaluation" (running title) has been initiated;
- At least 6 scientific sessions and 3 short courses will be offered at various scientific conferences by our working group. These include
  - Two (2) short courses, a presentation and a "Content Hub" at the 2019 DIA annual meeting in San Diego, focusing on topics of Safety evaluation and Interactive Graphics;
  - Two (2) topic-contributed sessions at the upcoming 2019 Joint Statistical Meeting (JSM) in Denver (JSM attendees: Please attend session #557 on Data Monitoring Committee, Session #166 on Interactive Safety Graphics);

- An invited session on "Visualization of Clinical Trial Data" at the 2019 Midwest Biopharmaceutical Statistics Workshop; and

- One (1) short course & 2 invited sessions at the 2019 ASA BIOP industry regulatory workshop.

Through these various multidisciplinary efforts, the ASA BIOP-SWG is changing how drug safety and benefit-risk are incorporated into clinical development programs, which includes assessment, monitoring and reporting as part of the regulatory submission and review, and also planned in post-marketing surveillance. This altogether will bring into fruition a coherent drug safety lifecycle, based on sound clinical judgment and statistical rigor.

The ASA Biopharmaceutical section Safety Working Group website will have updated information about the working group activities.

References

1. World Health Organization. The importance of pharmacovigilance—Safety monitoring of medicinal products. World Health Organization, Geneva, Switzerland, 2002.

2. O'Neill, R.T. (2002). Regulatory Perspectives on Data Monitoring. Statistics in Medicine 21: 2831-2842.

3. Wang W., Whalen E, Munsaka M, Li J, Fries M, Kracht K, Sanchez-Kam M, Singh K & Zhou K (2017), "On quantitative methods for clinical safety monitoring in drug development", Statistics in Biopharmaceutical Research, DOI: 10.1080/19466315.2017.1409134, posted online: 19 Dec 2017 *https://doi.org/10.1080/19466315.2017.1409134.*

4. Izem R, Sanchez-Kam M, Ma H & Zink RC. Sources of safety data and statistical strategies for design and analysis: Post-market surveillance. Therapeutic Innovation & Regulatory Science, vol. 52, 2: pp. 159-169, First Published January 8, 2018.

5. Ma H, Russek-Cohen E, Izem R, Marchenko & Jiang Q. Sources of safety data and statistical strategies for design and analysis: Transforming data into evidence. Therapeutic Innovation & Regulatory Science, vol. 52, 2: pp. 187-198, First Published March 7, 2018.

6. Marchenko O, Jiang Q, Chuang-Stein C, Mehta C, Levenson M, Russek-Cohen E, Liu L, Sanchez-Kam M, Zink R, Ke C, Ma H, Maca J, Park S. Statistical Considerations for Cardiovascular Outcome Trials in Patients with Type 2 Diabetes Mellitus. Statistics in Biopharmaceutical Research, Printed on-line on February 15, 2017: *http://amstat.tandfonline.com/doi/full/10.1080/19466315.2017.1280411.*

7. Marchenko O, Russek-Cohen E, Levenson M, Zink RC, Krukas M & Jiang Q. Sources of safety data and statistical strategies for design and analysis: Real world insights. Therapeutic Innovation & Regulatory Science, vol. 52, 2: pp. 170-186, First Published November 16, 2017.

8. Zink RC, Marchenko O, Sanchez-Kam M, Izem R, Ma H & Jiang Q. Sources of safety data and statistical strategies for design and analysis: Clinical Trials. Therapeutic Innovation & Regulatory Science, vol. 52, 2: pp. 141-158, First Published November 12, 2017.

9. Seltzer J, Li J and Wang W (2019), "Interdisciplinary Safety Evaluation and Quantitative Safety Monitoring: Introduction to a Series of Papers ", Therapeutic Innovation & Regulatory Science, Mar 29, 2019, *https://doi.org/10.1177/2168479018793130.*

10. Ball G, Kurek R, Hendrickson B, Buchanan J, Wang W W, Duke S, Bhattacharyya A, Li M, O'Brien D, Weigel J, Wang W Q, Jiang Q, Ahmad F, Seltzer J, Herrero-Martinez E, Tremmel L (2019), "Global Regulatory Landscape for Aggregate Safety Assessments: Recent Developments and Future Directions", Therapeutic Innovation & Regulatory Science, 1-5, DOI: 10.1177/2168479019841314, May 12 , 2019, *https://journals.sagepub.com/doi/abs/10.1177/2168479019841314 .*

11. Colopy M, Gordon R, Ahmed F, Wang W, Duke S (2018), "Statistical Practices of Safety Monitoring: An Industry Survey", Therapeutic Innovation & Regulatory Science, I-8, DOI: 10.1177/2168479018779973. ■

ASA BIOP REAL-WORLD EVIDENCE SCIENTIFIC WORKING GROUP UPDATE

Martin Ho, FDA and Weili He, Abbvie

FDA released its Framework for Real-World Evidence (RWE) Program in December 2018. It defines RWE as Clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of real-world data (RWD). RWD relates to patient health status and/or the delivery of health care routinely collected from a variety of sources, and RWE refers to evidence derived from RWD through the application of research methods. Due to new technologies and routine collection of electronic data in health claims data, medical records, and other clinical and administrative data, the real-world data have become more readily available.

The real-world data are highly pragmatic in nature and thus, the RWE is expected to be more generalizable than the evidence found in traditional randomized clinical trials. The U.S. Food and Drug Administration (FDA) PDUFA VI called for enhancing the use of RWE in regulatory decision making and committed public workshops and draft guidance between 2018 and 2021. The FDA has released multiple guidance and draft guidance documents related to RWE (FDA CBER/CDER 2013, 2018b; FDA CBER/CDER/CDRH 2013, 2017, 2018a; FDA CBER/CDRH 2017b; FDA CDER 2017a; FDA CDER/CBER 2018c). EMA also has published multiple guidance documents on the usage of RWE (e.g. the 2017 EMA Patient Registry Initiative to support research in understanding natural history of disease and characterizing the effectiveness and safety of products). Therefore, these new data-generating sources can help fill the gaps of the current expensive and suboptimal clinical trial enterprise.

Yet, there are many challenges of using RWD and RWE such as data standard, quality, relevant study design, and evidence synthesis methodologies that minimize potential biases and confounding inherent with RWD. Consensus has yet to be reached within the statistical community on how to measure the potential bias and level of uncertainty of the new class of synthesized evidence. As quantitative scientists, statisticians are well-equipped to take the lead in addressing these challenges translating these RW data into evidence to inform regulatory decisions.

With the anticipated greater uptake and utilization of RWD/RWE in medical research, we developed and submitted a proposal of a Scientific Working Group on RWE to the ASA BIOP Executive Committee in October 2017. We envision a group of statisticians from industry, academic, and regulator working together closely in a pre-competitive manner for transparency and advancement of statistical science. The group has a focus on the statistical aspects of RWD and RWE research and utilization, complementing ongoing research efforts in the broader RWE scientific community. The two primary goals of the group are (a) to advance understanding of the current landscape of RWE research and engage regulators in providing guidance and/or guiding principles on RWE research, and (b) to facilitate utilization and implementation of RWD and RWE using innovative and fit-for-purpose statistical methods in clinical research and medical product life cycle. Thanks to the Committee's helpful and enthusiastic support resulting in approval of the group's charter, the ASA RWE Scientific Working Group (RWE SWG) has formally been established in March 2018.

After the first in-person kickoff meeting at JSM 2018, the WG established consensus on the WG's scope of work, potential topics of interest, and potential deliverables. The WG has decided to adapt the "divide-and-conquer" approach and grouped the tentative topics of interests into two groups. This approach not only allows both groups proceed in parallel and save time, but it also encourages engagement of WG members in activities conducted in parallel to best leverage their diverse expertise and interests. Weili and Martin consulted members and decided the two workstreams would be divided by the regulatory use of RWD/RWE: Workstream #1 focuses on using RWD/RWE to support label expansions for medical products that are already on the market while workstream #2 focuses on using RWD/RWE to inform clinical study design. After considering individual members' expertise and preferences, the WG have decided on the membership and co-leads of both workstreams. Members can choose to participate in more than one workstreams. Table 1 lists our current members.

Table 1. List of RWE SWG Members by Affiliations

Industry	Affiliation	Academic/FDA†	Affiliation
Weili He [§,1]	AbbVie	Martin Ho [§,2]	CBER
Jie Chen	Merck	Telba Irony	CBER
Yixin Fang	AbbVie	Mark van der Laan	UC Berkeley
Douglas Faries	Eli Lilly and Company	Hana Lee	CDER
Qi Jiang	Seattle Genetics	Mark Levenson [1]	CDER
Kwan Lee [2]	Janssen	Zhaoling Meng	BMGMR[‡]
Xiwu Lin	Janssen	Pallavi Mishra-Kalyani	CDER
Yang Sung	Vertex Pharma. Inc.	Frank Rockhold	Duke
Hongwei Wang	AbbVie	Tingting Zhou	CBER
Roseann White	The Third Opinion	Ben Goldstein	Duke
Richard Zink	Target Pharma. Solution		

§ Co-chairs of the SWG
1 Co-leads of Workstream #1
2 Co-leads of Workstream #2
† Liz Stuart (Johns Hopkins University) participates as non-member
‡ Bill & Melinda Gates Medical Research Institute

Both workstreams follow a similar process to achieve their goals. First, each workstream identified 3-5 main topics. Next, workstream members were divided into small teams organically to work on identified key topics (see Table 2).

Table 2. List of identified key topics by ASA RWE SWG Workstream

Workstream #1: Use RWD/RWE for label expansion	Workstream #2: Use RWD/RWE to inform study design & analysis
• Regulatory, scientific, and ethical issues • Data sources, study types, and outcome measures • Estimands (treatment effect) in RW setting • Control of confounding	• Study of retrospective data only • Prospective study with external control or borrowing (Bayesian and frequentist approaches) • Causal inference framework in regulatory setting

Each team conducted a focused literature review to address the following four questions for a given topic: its regulatory context, a precise problem statement, a summary of current approaches, and gap analysis. Each team took turns to report their findings during monthly tele-conference calls. After receiving feedback from other group members, the teams would incorporate them into a write-up. Finally, the workstream co-leads will combine these sections and harmonize them into manuscripts, with a goal to publish them in a peer-reviewed journal.

After only eight months, the group has gone over these topics and met the milestones ahead of time. We are pleased to report that we presented a summary of preliminary findings of both workstreams in this year's Society of Clinical Trials Annual Meeting and in an invited session on RWE at JSM. Both were well received with strong interest from the attendees. In addition, the WG will also present their work to date in an RWE session at the upcoming ASA BIOP RISW in September 2019. The group is currently drafting manuscripts.

We want to share a few interesting preliminary observations with you. Firstly, we appreciate the importance of understanding data heterogeneity and fit-for-purpose approach in evaluating data quality. "Missing data" can be a misleading notion when describing gaps in availability of RWD: Some absent data entries are not "missing" because they are not supposed to be collected for its own primary purpose of use. For example, no claims data would be collected for those who drop out of a given insurance plan and it is by no means "missing." Secondly, gaps between historical RCT data and RWD may never close completely because their generation and report purposes differ. Rather, targeting selected data fields may be a more pragmatic approach. Thirdly, like in all rigorous and valid clinical studies, formulating estimands appropriately is crucial but it carries its unique set of challenges in the RW setting.

Finally, treatment modeling approach coupled with outcome masking (Li et al. 2016; Lim et al. 2018; Xu et al. 2019; Yue et al. 2014) is thus far the only causal inference study design used in regulatory submissions. However, potentially limited overlapping between comparative treatment groups remains a challenge, e.g., a large and diverse pool of patients as external controls may not be available for rare disease. Moreover, it has been a general challenge to predict the extent of overlapping between patients in the control pool and yet-to-be-enrolled patients in the prospective single-arm study. Since level of such overlap determines the power of a single-arm with external control study, the sample size estimation conducted at the design stage greatly relies on simulations for various scenarios and can be highly variable. (Li et al. 2016)Therefore, we are exploring the possibility of using doubly robust approaches that have been used extensively in areas outside regulatory submissions in last decade (Bang and Robins 2005; Funk et al. 2011; Kennedy et al. 2017; Laan et al. 2003; Pirracchio et al. 2015; Rubin and Stuart 2005). Since these approaches would involve both modeling treatments and outcomes, additional rigorous steps to mitigate potential Type I error rate would be paramount.

To conclude, we would like to invite all of you to attend our presentation at this year's ASA BIOP Regulatory-Industry Statistics Workshop. We look forward to sharing what the WG has done so far and to hearing your thoughts. See you there!

We want to thank the editor of the Report, Dr. Ilya Lipkovich from Eli Lilly and Company for his excellent comments and suggestions.

References

1. Bang, H., and Robins, J. M. (2005), "Doubly Robust Estimation in Missing Data and Causal Inference Models," Biometrics, 61, 962–973. *https://doi.org/10.1111/j.1541-0420.2005.00377.x*.

2. FDA CBER/CDER (2013), "Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets," U.S. Food and Drug Administration, Available at *http://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic.*

3. FDA CBER/CDER (2018b), "Use of Electronic Health Record Data in Clinical Investigations Guidance for Industry," U.S. Food and Drug Administration, Available at *http://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry.*

4. FDA CBER/CDER/CDRH (2013), "Electronic Source Data in Clinical Investigations," U.S. Food and Drug Administration, Available at *http://www.fda.gov/regulatory-information/search-fda-guidance-documents/electronic-source-data-clinical-investigations*.

5. FDA CBER/CDER/CDRH (2017), "Use of Electronic Records and Electronic Signatures in Clinical Investigations Under 21 CFR Part 11 –," U.S. Food and Drug Administration, Available at *http://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-records-and-electronic-signatures-clinical-investigations-under-21-cfr-part-11.*

6. FDA CBER/CDER/CDRH (2018a), "Use of Electronic Health Record Data in Clinical Investigations Guidance for Industry," U.S. Food and Drug Administration, Available at *http://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry.*

7. FDA CBER/CDRH (2017b), "Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices," U.S. Food and Drug Administration, Available at *http://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices*.

8. FDA CDER (2017a), "E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials," Available at */regulatory-information/search-fda-guidance-documents/e9r1-statistical-principles-clinical-trials-addendum-estimands-and-sensitivity-analysis-clinical*.

9. FDA CDER/CBER (2018c), "Framework for a Real-World Evidence Program," Federal Register, Available at *https://www.federalregister.gov/documents/2018/12/07/2018-26546/framework-for-a-real-world-evidence-program-availability*.

10. Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., and Davidian, M. (2011), "Doubly Robust Estimation of Causal Effects," *American Journal of Epidemiology,* 173, 761–767. *https://doi.org/10.1093/aje/kwq439*.

11. Kennedy, E. H., Ma, Z., McHugh, M. D., and Small, D. S. (2017), "Non-parametric methods for doubly robust estimation of continuous treatment effects," *Journal of the Royal Statistical Society:* Series B (Statistical Methodology), 79, 1229–1245. *https://doi.org/10.1111/rssb.12212*.

12. Laan, M. J. van der, Laan, M. J., and Robins, J. M. (2003), Unified Methods for Censored Longitudinal Data and Causality, Springer Science & Business Media.

13. Li, H., Mukhi, V., Lu, N., Xu, Y.-L., and Yue, L. Q. (2016), "A Note on Good Practice of Objective Propensity Score Design for Premarket Nonrandomized Medical Device Studies with an Example," *Statistics in Biopharmaceutical Research*, 8, 282–286. *https://doi.org/10.1080/19466315.2016.1148071*.

14. Lim, J., Walley, R., Yuan, J., Liu, J., Dabral, A., Best, N., Grieve, A., Hampson, L., Wolfram, J., Woodward, P., Yong, F., Zhang, X., and Bowen, E. (2018), "Minimizing Patient Burden Through the Use of Historical Subject-Level Data in Innovative Confirmatory Clinical Trials: Review of Methods and Opportunities," *Therapeutic Innovation & Regulatory Science*, 52, 546–559. *https://doi.org/10.1177/2168479018778282.*

15. Pirracchio, R., Petersen, M. L., and van der Laan, M. (2015), "Improving Propensity Score Estimators' Robustness to Model Misspecification Using Super Learner," *American Journal of Epidemiology,* 181, 108–119. *https://doi.org/10.1093/aje/kwu253*.

16. Rubin, D. B., and Stuart, E. A. (2005), "Matching with Multiple Control Groups, and Adjusting for Group Differences," 12.

17. Xu, Y., Lu, N., Yue, L., and Tiwari, R. (2019), "A Study Design for Augmenting the Control Group in a Randomized Controlled Trial: A Quality Process for Interaction Among Stakeholders," *Therapeutic Innovation & Regulatory Science,* 2168479019830385. *https://doi.org/10.1177/2168479019830385*.

18. Yue, L. Q., Lu, N., and Xu, Y. (2014), "Designing Premarket Observational Comparative Studies Using Existing Data as Controls: Challenges and Opportunities," *Journal of Biopharmaceutical Statistics,* 24, 994–1010. *https://doi.org/10.1080/10543406.2014.926367* . ■

6TH NONCLINICAL BIOSTATISTICS CONFERENCE: OVERVIEW

Steven Novick and John Kolassa, conference co-chairs

One hundred and thirty attendees converged on the Rutgers University campus Student Center in New Brunswick, New Jersey, to participate in the 2019 Nonclinical Biostatistics Conference from June 17 - 19. Held for the sixth time, the biennial conference provided a venue for the presentation and discussion of scientific and statistical issues relevant to the field of nonclinical biostatistics. Jointly organized by the Biopharmaceutical Section of the ASA in collaboration with the Department of Statistics at Rutgers University, the program featured 29 invited and contributed talks, 26 poster presenters, special sessions for graduate students, and was highlighted by keynote addresses, delivered by Dr. Karen Kafadar (ASA President-elect) and Dr. José Pinheiro (Global head of statistical modeling & methodology, Janssen)

To kick off the conference, two short courses were offered:

An R shiny tutorial with nonclinical applications *Instructors: Max Kuhn and Phil Bowsher, RStudio*

Getting it right: Compositional analysis of biological measurements *Instructors: Anthony Lonardo (Lonardo StatReg Associates) and Juan José Egozcue and Maribel Ortego (Dept. Civil and Environmental Engineering, Universitat Politecnica de Catalunya in Barcelona, Spain)*

The Best Nonclinical Biostatistics Paper Award was presented to the authors of:

Zeng, L., Novick, S., Yu, B., and Yang, H. (2019). "General Framework for Equivalence Testing over a Range of Linear Outcomes with CMC Applications", Statistics in Biopharmaceutical Research, 11(2), 182-190.

Two graduate students walked away with best poster award with first prize ($250) going to Perceval Sondag and second prize ($150) going to Yi Hua.

All of the 2019 NCB conference oral presentations and posters are available electronically. Please visit our website at *http://community.amstat.org/biop/events/ncb/index*.

We thank all who contributed to the success of the conference, including our sponsors, organizers, presenters, attendees, and helpers. We hope to see you next time in 2021!

Questions, suggestions and comments may be directed to *Xin.Huang@abbvie.com* or to *NovickS@medimmune.com*

LinkedIn with NCB at *www.linkedin.com/groups/8547808.* ■

THE 2019 ASA BIOPHARMACEUTICAL SECTION REGULATORY-INDUSTRY WORKSHOP

Judy Li , Celgene and Renée Rees, FDA/CBER

On behalf of the Steering Committee, we are pleased to welcome the attendants of the 2019 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop on September 23-25 at the Marriott Wardman Park Hotel in Washington DC. The theme for this year's Workshop is "From Small Data to Big Data, from RCT to RWE, the Impact of Statistics". This year's program includes

- 2 plenary sessions
- 42 parallel sessions
- 10 short courses
- 42 roundtable discussions
- 40 posters

In addition, a mixer on September 24 (Tuesday) evening will provide more opportunities for networking and socializing. Detailed information on the workshop can be found via the link at *https://ww2.amstat.org/meetings/biop/2019/program.cfm*.

Compared to previous workshops, this year we are offering an increased number of short courses, posters and student travel awards. We would also like to highlight some exciting new features this year. For the first time, we are soliciting sponsor support to help keep registration fees affordable and provide a better workshop experience for attendees. With that, we are adding a small souvenir for all attendees, which will come in the workshop tote. To attract a larger student attendance, we have also granted a few student registration waivers. In addition to the new features, this year's workshop will continue to offer popular aspects: meeting app, poster awards, mixer and enhanced audio and visual support. The plenary sessions will be closely related to the theme of the workshop and eminent members from FDA, industry and academia will participate in both plenary sessions. The first plenary session features two keynote speakers. Dr. Jacqueline Corrigan-Curay from FDA CDER's Office of Medical Policy will discuss the regulatory aspects of the potential use of real world evidence (RWE) in regulatory decisions by reviewing the ongoing FDA RWE program. The title of her talk is "FDA RWE Program -Informing Regulatory Policy". The secondary plenary keynote speaker, Dr. Pandu Kulkarni from Eli Lilly and Company will showcase recent development and application of advanced analytics methodologies for innovative designs and analyses within the health care industry, such as artificial intelligence (AI) and machine learning (ML). The title of his talk is "Leading the Future of Health Care Industry with Advanced Analytics, Artificial Intelligence, and Machine Learning". The secondary plenary session start with one keynote speaker, Professor Mark van der Laan from UC Berkeley. Dr. van der Laan will discuss a general roadmap for generating causal inference based on observational studies used to generate RWE. The title of his talk is "Targeted Machine Learning for Causal Inference based on Real World Data". A panel discussion will follow, which will include top experts Dr. Aloka Chakravarty, FDA-CDER, Dr. John Scott, FDA-CBER, Dr. Ram Tiwari, FDA-CDRH, Dr. Michael Branson, Celgene Co., William Wang, Merck & Co. along with Dr. van der Laan.

The workshop also features ten short courses, including state of art and contemporary topics in modern statistical methodologies and clinical development. All short courses will take place on Monday, September 23. The ten short courses present the following topics:

- Simulation Practices for Adaptive Clinical Trial Design in Drug and Device Development
- Biomarker-Assisted Clinical Designs: Concepts, Rationale, and Case Studies
- Designing and Integrating RCT/RWE in Safety Decision-Making
- Methods for Causal Inference from Randomized Trials with Loss to Follow-Up or Non-Adherence
- Statistical Analysis of Composite Endpoints in Clinical Trials

- Smart Simulation with SAS and R
- New Adaptive Design Guidance
- Flexible Sample Size Designs with Applications to Improve the Efficiency and Probability of Success of Industry-Sponsored Clinical Trials
- Real-World Data and Evidence: An Interdisciplinary Approach and Applications to Precision Medicine and Health Care
- Use of Historical Data in Clinical Trial: An Evidence Synthesis Approach

The 42 parallel sessions and 42 round-table luncheon sessions cover a wide range of current and hot topics regarding the use of the cutting-edge statistical science in health care; some of these sessions will surely match your interests!

We would like to express our sincerest gratitude to the many people involved in the workshop planning and execution. This includes but is not limited to the ASA Biopharmaceutical Section Executive Committee for their guidance and support; ASA meeting planning support, especially Ms. Kristin Mohebbi; and all the Workshop Steering Committee members and advisors. We would also like to thank all the participants. Without their contribution, the workshop would not be successful. ■