Volume 29, No. 1 • Spring 2022

BIOPHARMACEUTICAL REPORT

Chair: Alan Hartford **Editors:** Herbert Pang, Ling Wang, Kristi L. Griffiths

CONTENTS

Notes from the editors

Hope this message finds you well and we are all very glad to see that COVID-19 is under better control this Spring. It is now easier for us to reconnect with colleagues, family, and friends in person. One of the lessons learned from this pandemic is that digital tools and technologies present significant opportunities for us. Our first issue of 2022 will focus on the theme of "digital health" with featured contributing articles from industry, government, and academia .

In this first issue of 2022, we open with an article from our current Section Chair, **Alan Hartford,** and the 2021 BIOP chair, **Weili He**, reflecting on the past year and sharing the plans for the rest of 2022. Next, we feature an article by **Junrui Di** (Pfizer), **Jiawei Bai** (Pfizer), **Isik Karahanoglu** (Pfizer), **Nunzio Camerlingo** (Pfizer), and **Charmaine Demanuele** (Pfizer). They describe the analytical challenges of multi-modal sensor data and discuss the practical considerations for incorporating them into clinical trials. This is followed by an article by academics and NIH's National Institute of Mental Health (NIMH). **Vadim Zipunnikov** (Johns Hopkins), **Debangan Dey** (Johns Hopkins), **Kathleen Merikangas** (NIMH), and **Andrew Leroux** (U Colorado, Denver) present the statistical challenges of modelling mobile digital health data and advocate for a collaborative team effort to translate the research effort to actionable health information. Next up is a feature article contribution from non-clinical statistics colleagues, **Elliott Schmitt** (Moderna), **Gang Wang** (Moderna), and **Julia O'Neill** (Moderna), on the use of Digital Twins for accelerating process development. They also present a case study focusing on the model development and application aspects. Our fourth feature article is by **Irina Gaynanova** (Texas A&M). She outlines reproducibility challenges associated with continuous glucose monitoring data as digital biomarkers of glucose control, and highlights the accompanying free and open-source R package and Shiny app *iglu* for the calculation of various continuous glucose monitoring metrics. This is followed by a second non-clinical statistics article. **David Christopher** (Merck), **Erik Talens** (Merck), and **Phillip Yates** (BMS), leaders from the Chemistry, Manufacturing, and Controls Statistics Leadership Group and Biostatistics Statistics Leadership Group of the International Consortium for Innovation and Quality in Pharmaceutical Development (IQ Consortium), introduce IQ leadership forums and promote IQ as a unique platform for cross-industry collaboration. Later in this issue, you will find a summary report from a virtual discussion organized by the ASA BIOP Statistical Methods in Oncology Scientific Working Group, the FDA Oncology Center of Excellence, and LUNGevity Foundation. The topic of discussion is "Statistical Considerations in Clinical Trials for Rare Pediatric Cancers". The final article is from **Margaret Gamalo** (Pfizer), Editor-in-Chief for the *Journal of Biopharmaceutical Statistics*. She shares a summary of *JBS*'s special issues published in 2022 and plans of future issues. We would like to continue this series by inviting other editors of statistical journals with broad readership among the BIOP membership. In the last section, we provide an update on upcoming conferences in 2022 that are of interest to the BIOP community. The editors would like to thank all the authors of the articles for their time and contributions, and wish that everyone enjoys this very first issue of the BIOP Report in 2022.

TRANSITION REPORT

Weili He (BIOP Chair 2021) and Alan Hartford (BIOP Chair 2022)

Greetings to all BIOP Section members. We want to thank you for the honor of serving the members of the Biopharmaceutical Section (BIOP). 2021 was another extraordinary year, a time of both tragedy and exciting breakthroughs. It continued to be a unique year, while we perfected the way of working in a pandemic era and still managed to successfully achieve our goals of serving the BIOP community. To no one's surprise, but to all their credit, our BIOP committees have been successfully operating virtually and continue to be extremely effective in providing benefits to our members and preparing for our annual meetings. But there were many changes in the BIOP as there were everywhere.

As we continued to work virtually in 2021, all 3 BIOP EC meetings were held virtually. While we all missed the in-person interactions, the EC still moved forward with section business without any hiccups. Also held virtually was the annual business meeting during JSM. In the pre-pandemic years, this was a great event for networking, catching up with acquaintances, learning about section business and seeing the award winners while having some food and drink supplied by BIOP. As with 2020, the 2021 business meeting was virtual; Nonetheless, there were approximately 90 attendees for the annual update to the BIOP business.

Our annual Regulatory-Industry Statistics Workshop, even still held virtually, was another huge success where the virtual attendance reached 1300. The interesting and contemporary topics and the smooth running of the workshop without many technical glitches afforded the attendees a pleasant learning and networking experience. The ASA, workshop co-chairs, and the organizing committee did an amazing job putting together an excellence workshop program.

In this article, we provide a summary of the most important initiatives and events that took place in 2021 and early 2022, and our plans for the rest of 2022.

Biopharmaceutical Section Scholarship

The Biopharmaceutical Section scholarships were awarded again in 2021. Consideration for the awards is based primarily on notable academic achievement or applied project work related to the area of biopharmaceutical statistics. General academic performance, leadership, volunteering, and service will also be reviewed. The 2021 recipients were:

- **Siddhesh Kulkarni** (University of Connecticut)
- **Inkoo Lee** (Florida State University)
- **Ruizhe Chen** (University of Illinois at Chicago)
- **Michael Daniel Lucagbo** (University of Maryland, Baltimore County)

Student Paper Award at Joint Statistical Meetings

First: **Tian Gu,** U of Michigan on *An ensemble meta-prediction framework to integrate multiple external models*

Second: **Bingkai Wang,** Johns Hopkins on *Precision by Stratified Randomization and Covariate Adjustment*

Third: **Sharon Ling,** U of Minnesota on *Calibrated Dynamic Borrowing Using Capping Priors*

Honorable Mention: **Nathan Bean,** U of North Carolina on *Bayesian Multi-Regional Clinical Trials Using Model Averaging*

Honorable Mention: **Ethan Alt,** U of North Carolina on *Historical Data with Strict Control of Family-wise Error Rate*

Biopharmaceutical Section 40th Anniversary

The Biopharmaceutical Section celebrated its 40th anniversary as a section in 2021. The committee planned a panel session with past BIOP chairs at the 2021 Joint Statistical Meeting. There were also several articles included in 2021 Biopharmaceutical reports (Spring, Summer, Fall & Winter issues) on the BIOP section

journey over the years and reflection from several past BIOP section chairs on their experience.

Due to the virtual nature of both JSM and RISW in 2021, BIOP EC also has planned on its 40th +1 anniversary celebration at both 2022 JSM and RISW, if held in-person. Stay tuned!

Joint Statistical Meetings

The Section had a successful presence at the virtual JSM in 2021. Special thanks and acknowledgment to Jonathan Moscovici, who as our BIOP 2021 Section Program Chair, for putting together a spectacular program. Among the 12 invited session proposals submitted to the section, the whole JSM invited program included five Biopharmaceutical Section sponsored sessions. Topics included stories from COVID-19 vaccine development: Statistical challenges and opportunities, statistical challenges linked to estimands of interest, synthetic clinical trial design to accelerated FDA approvals, data science and statistics in pharmaceutical engineering and experimental studies: where is it, where is it going, and considerations in clinical trials endpoints selection. There were 14 topic contributed sessions sponsored by BIOP, selected from 33 submitted session proposals. Also included were 118 contributed abstracts for 6 speed sessions and 9 roundtable sessions. This is another stellar year for the BIOP at JSM.

ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop

As noted above, 2021 was another record setting year for attendance at the Regulatory-Industry Statistics Workshop under the leadership of co-chairs Gene Pennello and Bo Huang and their steering committee. However, looking forward, thanks to the efforts from ASA meeting planners and the BIOP Workshop Task Force, when we return to an in-person venue for the Workshop, hopefully in 2022, we will have a new space at the Bethesda North Marriott Hotel and Conference Center which will allow for growth and additional flexibility for activities within an in-person Workshop.

We are also very happy to report that the 2021 Workshop sponsorship was highly successful! There were 8 principal sponsors and 8 supporting sponsors, raising over $54,000 to support Workshop activities. Sponsorships for the 2022 Workshop are already underway, with numerous opportunities remaining. For more information on Workshop sponsorships, please see *https://ww2.amstat.org/meetings/biop/2022/sponsors.cfm*. We look forward to the exciting program that our 2022 co-chairs Chia-Wen (Kiki) Ko and Hope Knuckles and their steering committee will provide for the attendees.

ASA Biopharmaceutical Section Nonclinical Biostatistics Conference

The 2021 Nonclinical Biostatistics Conference was held virtually June 21-24, 2021. The conference was the 7th such conference since 2009, meeting biennially. It was organized by the ASA BIOP section's nonclinical working group, co-chaired by Xin Huang (Abbvie) and John Kolassa (Rutgers). One-hundred forty attendees participated in a program that kicked off with two short courses: Bayesian Regression Trees BY Dr. Jason Roy (Rutgers) and Bayesian Survival and Joint Models using Rstanarm, Jacqueline Buros Novik (Generable Inc.). There were 28 technical presentations and 14 posters related to the 4 main areas of nonclinical biostatistics (Discovery/Biomarkers, Safety/Pharmacology, CMC, Statistical Computing and Visualization). The conference also recognized 3 awardees for the best nonclinical papers published over the preceding 3 years as follows:

- **1st Place: Burdick, R. K., Thomas, N., & Cheng, A. (2017)**. Statistical considerations in demonstrating CMC analytical similarity for a biosimilar product. Statistics in Biopharmaceutical Research, 9(3), 249-257.
- **2nd Place: Novick, S. J., Christian, E., Farmer, E., & Tejada, M. (2021)**. A Bayesian statistical approach to continuous qualification of a bioassay. PDA Journal of Pharmaceutical Science and Technology, 75(1), 8-23.

- **3rd Place: Sondag, P., & Lebrun, P. (2020).** Risk-based similarity testing for potency assays using MCMC simulations. Statistics in Biopharmaceutical Research, 1-10.

Two graduate students received best poster awards, with first prize ($250) going to Louise Leonard and second prize ($150) going to Jinghang Lin.

Fellows Committee

Our BIOP Fellows Committee will identify potential fellows from within BIOP and provide guidance in submitting nomination packages. Please utilize this resource if you are considering application for an ASA Fellow or sponsoring someone. A wonderful article was produced by BIOP members in 2020 about Advice for ASA Fellow Nominations. If you haven't read it, check it out here: *https://magazine.amstat.org/blog/2020/06/01/advice-for-asa-fellow-nomination/*

Scientific Working Groups

The process of establishing new working groups is overseen by the Scientific Working Group (SWG) Committee. Through the SWG Committee, section members can submit research topics that contribute to the goals of advancing the science, enabling innovation, and leveraging the membership expertise. The establishment of the working group must be approved by the Section Executive Committee and each scientific working group must provide a yearly update report to the Executive Committee.

Individuals interested in forming a new SWG can review the BIOP guidelines for more information. *https://community.amstat.org/biop/aboutus/sub-committees/swg*

Outreach

We are always looking for new ways to build synergies with other groups and share best practices within our own membership!

For 2021, the BIOP EC recognized the importance of our Outreach and Collaboration Committee. Our BIOP Manual of Operations was updated to specify that the BIOP Chair-Elect will become a member of the Outreach Committee to ensure ongoing engagement with other Sections and Chapters as well as external organizations.

The BIOP is very fortunate to have generous sponsors and a large membership, which means we have available funds to benefit our members' section experiences.

BIOP was very active in 2021 on our publication front, with 11 new podcasts from Richard Zink, 9 BIOP sponsored webinars, 4 issues of the Biopharmaceutical Report at *https://community.amstat.org/biop/biopharmreport*.

We completed 7 years of the BIOP mentoring program! This program has been beneficial in pairing students and young professionals with more veteran members of the Section to provide career advice. We are looking forward to many more years of this successful program. For more information on the BIOP mentoring program, send an email to *BiopharmMentoring@gmail.com*. The BIOP mentoring program was also a featured topic in the 2020 Podcast series. Check out podcast Episode 81 to hear from 3 mentor-mentee pairs on the value of the mentoring program.

All the Podcast Episodes can be found here: *https://www.buzzsprout.com/16296*

While BIOP is one of the largest ASA Sections, we have made concerted efforts to highlight the benefits of BIOP membership at our conferences and in our publications (see Vol 26, issue 2 of the Biopharmaceutical Report). We believe that we provide a lot of educational and networking opportunities to our members. We hope you think so too. The best recommendations we can receive are from our members, so we would appreciate your recommendations for people to join the Section!

Membership Survey

We conducted a membership survey in early 2021, and you can find the survey results in the BIOP report Vol 28 Issue 2. Based on the feedback, we will continue to strive to improve our service to the members.

Communication

We have revised and created roles for publications/communication, including adding Communication

Secretary and Social Media Coordinator in addition to the elected Publication Officer. We are examining new platforms to get the information out, and to hear from our members. Stay tuned in 2022 for more on this topic.

Biopharmaceutical Section Newly Elected Officers

We would like to welcome the following elected officers to the Biopharmaceutical Section Executive Committee for 2021:

- Chair-Elect 2022: **Brian A. Millen**
- Program Chair-Elect 2022: **Elena Polverejan**
- Publications Office 2022-2024: **Hiya Banerjee**
- Council of Sections Representative 2022-2024: **Janelle Charles**

We know these individuals will do a fantastic job representing the Section and we wish them the best of luck!

2022 plan

For 2022, we have both new and renewed efforts underway. We have a new committee, Statisticians in Small Biotech, led by Liang Fang, to focus on (you guessed it!) serving the needs of statisticians working in smaller companies who may not have other statisticians on staff with whom to share information, both technical and procedural. This committee will focus on identifying challenges and opportunities that are unique to this group and will work on solutions and ideas for these challenges and opportunities, making our BIOP community even more valuable to them. These efforts will likely also be of value to all statisticians starting their career to augment the training they may receive from their employers.

One of our renewed efforts underway is to further our connections with statistical leaders across our section. Our goal is to explore the concept of leadership with our section members. With this in mind, please share your stories of examples of leadership you've seen of statisticians working across BIOP by sending an email to *asabiopharm@gmail.com*. We'd like to share your stories of how you've witnessed the impact they have made that has earned your admiration.

As mentioned above, we are putting additional focus on BIOP's Outreach and Collaboration Committee. This year we will review and update the charter of this committee. We will deliberate on how we can build a "bigger tent" for potential new members.

Final Thoughts

We would like to take this opportunity to thank all the elected officers, committee chairs, and committee members for their commitment, time, energy, and expertise in the smooth running of the Section. Without all of you, the Section would not have been able to accomplish everything that it did. Despite the pandemic, the dedication to the section did not waiver. That is greatly appreciated.

Special thanks to our outgoing elected officers at the end of 2021:

- **Bruce Binkowitz,** Past Chair (Chair in 2020)
- **Jonathan Moscovici,** Program Chair
- **Yongming Qu,** Publication Officer
- **Veronica Bubb,** Council of Sections Representative

Curious to see what it takes to keep the Section running? Check out our Charter and our Manual of Operations under the About Us tab of our website *https://community.amstat.org/biop/home*. Thanks to Bruce Binkowitz for leading the 2021 updates to the Manual of Operations.

Finally, we would like to thank the membership for your support of our very active section. We are looking forward to a productive 2022, and hopefully seeing you all in person again, in 2022! ■

Be Safe, Be Well,
Weili and Alan

DEPLOYMENT AND APPLICATION OF MULTI-MODAL SENSORS IN CLINICAL TRIALS

Junrui Di (Pfizer), Jiawei Bai (Pfizer), F. Isik Karahanoglu (Pfizer), Nunzio Camerlingo (Pfizer), Charmaine Demanuele (Pfizer)

1. Introduction

Over the last decade, Digital Health Technologies (DHTs) have proven to be effective in measuring human activity and physiology, and been integrated into numerous clinical trials (Digital Medicine Society (DiME), 2021). DHTs have evolved rapidly especially in recent years, with astonishing strides toward miniaturization and life-cycle extension. It has been reported previously (Karas *et al.*, 2019) that small sensors such as accelerometers could provide valuable information about the study participants' health condition. While more of these sensors (e.g. heart rate monitor, thermometer, pulse oximeter (SpO2), continuous glucose monitoring) were being packaged together in modern wearable devices, one of the most common questions to ask is, whether and how we can better understand human health using the data from all these sensors combined. This question arises naturally from the fact that each sensor often provides location-specific information, that is, limited to one dimension of human activity or physiology. For example, accelerometers measure the magnitude of physical movement of the body, while heart rate monitors may indicate the level of exertion. However, neither of these two sensors could reliably provide blood oxygen level, which would need an SpO2 sensor. For a study focusing on heart failure, each of the three domains (activity, exertion, and blood oxygen level) have its own clinical indication and missing any of them may fail to provide a holistic picture of disease progression. Therefore, combining multi-modal sensor data may enable more comprehensive phenotyping, better symptom characterization and more accurate assessment of changes in health status over time.

A few notable studies have explored this path. Merikangas *et al*. (2019) included both an accelerometer and an ecological momentary assessment (EMA) component in their study to examine the associations among motor activity, energy, mood, and sleep. The joint modeling of a) the motor activity and sleep measurement derived from accelerometry data and b) the mood and energy level assessed through the EMA devices offered the authors opportunities to gain insights into how these different domains interact with each other and potentially what the therapeutic target is for patients with bipolar disorder. A similar example is the Apple Women's Health Study (Mahalingaiah *et al*., 2021), which aims to investigate the relationship among women's menstrual cycles, health and behavior, through a "mobile-application-based longitudinal cohort study" that has both a sensor and a survey component. Their analysis aims to combine both the (monthly) survey data with longitudinally measured smartphone/watch data, so that a better understanding might be reached of how the menstrual cycle relates to exercise, sleep, environment, behavioral and other physiological processes. Besides typical observational studies, Quer *et al*. (2021) showcased that multi-modal sensor data including heart rate, sleep and activity coupled with self-reported symptoms could significantly distinguish between symptomatic individuals with and without a diagnosis of COVID-19

These studies all highlighted the fact that multi-modal sensors were beneficial because each sensor contributed distinct aspect of information to the statistical model. However, there are considerations researchers should be aware of, before conducting studies with multi-modal sensors. In the remainder of this article, we will first discuss typical analytical challenges, and then elaborate on the requirements of deploying multi-modal sensors in clinical studies.

2. Analytical Challenges of Multi-Modal Sensor Data and Emerging Techniques

Modern DHTs collect data passively, continuously, and frequently, leading to rich streams of time series data with high dimensionality, complex data structure, and potentially noisy signals. With features derived from multi-modal sensors, one can directly combine those features in linear or nonlinear fashion using statistical and machine learning models. For example, as discussed previously, by combining motion-related features acquired from actigraphy, and heart rate/heart rate variability features acquired from wearable electrocardiogram (ECG), the accuracy for sleep prediction and sleep stage classification can be potentially increased when compared to using only one of these modalities (Aktaruzzaman *et al*., 2017; Yuda *et al*., 2017). However, when multi-modal sensors are deployed simultaneously, some new challenges arise due to the continuous nature of the measurements, and the interrelation between different modalities. It becomes crucial to fully utilize the rich data and identify the homogenous underlying signals (such as disease progression or treatment effects) from multiple modalities while accounting for the possible heterogeneity across modalities.

2.1 Fully Utilize the Temporal Aspect of Sensor Data

Before fusing data collected by multiple sensor modalities, a key issue to consider is to leverage the continuous time series signals from each of the sensors. It is still common practice to derive features that quantify certain physiological or behavioral characteristics (Di *et al*., 2019). For example, total activity counts have been used to represent overall daily activity intensity in many studies using accelerometers (Varma *et al*., 2017). Similarly, time-in-ranges indices are commonly employed in studies involving continuous glucose monitoring (CGM) sensors to quantify the quality of glucose control (Battelino *et al*., 2019) However, these features are summary measures and do not reveal the temporal variations within a day.

In circadian rhythm research, cosinor and extended cosinor models have been utilized to parametrically estimate the daily diurnal trend as a cosinor (or transformed cosinor) curve to represent the amplitude and phase of time series data (i.e. time to reach the peak) (Marler *et al*., 2006; Cornelissen, 2014). Time series data collected by sensors can be considered as a function of time. More recently, functional data analysis (Georgiev *et al*., 1998), which was developed to study the smooth functional behaviors of curves over a continuum, has been widely use to nonparametrically estimate the temporal characteristics of physiological trends or diurnal patterns (Goldsmith *et al*., 2016). By assuming the underlying functional smoothness, functional data analysis approaches such as functional regression (function-on-scalar or scalar-on-function) and functional principal component analyses can identify treatment effects within a specific time window in a day, or to detect a shift of phase across different cohorts. For example, Spira *et al*. recently discovered significant differences in activity levels between participants with and without β-amyloid (Aβ) antibody only within specific time windows during a day, by using function-on-scalar regression (2021).

Other than emerging statistical methodologies that aims to reveal temporal trends, modern deep learning architectures such as Recurrent Neural Network (RNN) can ingest time-sequential data collected by wearables to solve for prediction problems (Nweke *et al*., 2018), such as human activity recognition (Chen *et al*., 2021). Specifically, RNN models using Long Short-Term Memory (LSTM) with different memory units have been widely used to model data collected by wearable devices (Rabby *et al*., 2021; Uddin and Soylu, 2021).

The prediction ahead of time of glucose concentration levels can be reliably achieved by exploiting their recent history, monitored by (minimally invasive or non-invasive) CGM sensors, in combination with data-driven algorithms. Simple data-driven strategies, using polynomial or linear autoregressive models (Eren-Oruklu *et al*., 2009), as well as more sophisticated methods, such as Kalman filters (Facchinetti *et al*., 2011) or neural networks (Rabby *et al*., 2021), have proven effective in the short-term prediction of future glucose levels (Prendin *et al*., 2021).

With the amount of available data collected by wearables rapidly growing, these deep learning

approaches and model-based techniques will become mainstream and standard approaches to deal with real-world measurements.

2.2 Separate the Joint Effects and Individual Modal Specific Effects

In clinical trials, fusion of multi-modal sensor data can be used to identify overall treatment effects by aggregating information from different physiological/behavioral domains. To reveal such effects, sometimes it is necessary to separate the joint effects that are homogenous across different modalities from the modal-specific effects.

In 2013, Lock *et al*., developed the Joint and Individual Variation Explained (JIVE) and used it to study the association between gene expression and miRNA data collected from the same samples (2013). As a data fusion technique and an extension to principal component analysis, JIVE decomposes multi-modal data into a low-rank approximation capturing joint variation across data types, low-rank approximations for structured variation individual to each data modal. Di *et al*. applied it to integrate accelerometry-derived features quantifying three physiological domains of activity, sleep, and circadian rhythm quantify and separate between- and within-domain variation (Di *et al*., 2019). The same concept can be directly implemented to study features derived from multi-modal sensors. With recent generalizations and extensions such as to account for heterogeneous data types (continuous/binary/count) (Li and Gaynanova, 2017) and partially shared information between modals (Gaynanova and Li, 2017), JIVE shows the promise to fuse multi-modal sensor data.

JIVE provides a framework to properly quantify the interrelation and codependency across multiple data modalities. Conceptually, the interrelation and codependency can be considered as an outcome measurement by itself. With longitudinal clinical trials with multi-modal sensors, the change of such interrelation can be traced and analyzed to provide meaningful clinical interpretation.

3. Practical Considerations to Incorporate multi-modal sensors into clinical trials

FDA recently released the draft guidance "Digital Health Technologies for Remote Data Acquisition in Clinical Investigations" which provided recommendations on the use of DHTs in clinical investigations, such as considerations for device selection, endpoints validation and verification, and statistical analysis. (US FDA, 2021). Di *et al*. provided operational suggestions to deploy DHTs in clinical studies to minimize the impact of missing data (Di *et al*., 2022). Incorporating multi-modal sensors should in principle follow these suggestions, such as to configure the devices appropriately, to determine the optimal placement location of the device, and to collect additional contextual information, when possible.

For clinical studies where patients wear one or multiple devices for a long period of time, it is crucial to incorporate the patients' perspective to increase their adherence. At the design phase of the studies, focus group of patients can be used to capture their voice to understand the preferrable form factor of the device(s) and the outcome measures that is the most meaningful to their daily life and health conditions. One question that can be considered is that to obtain a holistic view of multiple physiological/cognitive/behavior/environmental domains, should we identify a single device with multiple embedded sensors instead of providing multiple devices? To reduce risks to patients, as suggested by FDA in the draft guidance (US FDA, 2021), it is important to have a comprehensive informed consent of human subjects that details what data will be acquired from the multi-modal sensors, what foreseeable risks, patients' privacy concern, or discomforts may occur in using the sensors, and intended research purposes and data use.

There are other advantages of using multi-modal data to improve clinical studies. For example, with the technological advancement and widespread adoption of consumer grade wearable devices that contain multiple built-in sensors, we can obtain individualized baseline

that is close to "truth" using patients' historical wearable device data to assess change in digital measures over weeks, months, or even years. Similarly, this historical device data can help prescreen patients for particular phenotypes and characteristics of interest to select patient cohorts for early phase (I or II) studies. This has the potential to improve the efficiency of these typically small studies by reducing variability.

4. Conclusion

Multi-modal sensors are beneficial to clinical studies by providing a holistic picture of human behavior and physiology in real-life. A broader application of advanced methodologies and innovative approaches to analyze data from multi-modal sensors are needed for researchers to fully utilize those valuable data.

5. Reference

Aktaruzzaman, M. *et al*. (2017) 'Performance comparison between wrist and chest actigraphy in combination with heart rate variability for sleep classification', *Computers in Biology and Medicine*, 89, pp. 212–221. doi: 10.1016/j.compbiomed.2017.08.006.

Battelino, T. *et al*. (2019) 'Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range', *Diabetes Care,* 42(8), pp. 1593–1603. doi: 10.2337/dci19-0028.

Chen, L. *et al*. (2021) 'Deep learning based multimodal complex human activity recognition using wearable devices', *Applied Intelligence,* 51(6), pp. 4029–4042. doi: 10.1007/s10489-020-02005-7.

Cornelissen, G. (2014) 'Cosinor-based rhythmometry', *Theoretical Biology and Medical Modelling*. BioMed Central, 11(1), p. 16.

Di, J. *et al*. (2019) 'Joint and Individual Representation of Domains of Physical Activity, Sleep, and Circadian Rhythmicity', *Statistics in Biosciences,* 11(2), pp. 371–402. doi: 10.1007/s12561-019-09236-4.

Di, J. *et al*. (2022) 'Considerations to address missing data when deriving clinical trial endpoints from digital health technologies', *Contemporary Clinical Trials,* 113, p. 106661. doi: 10.1016/j.cct.2021.106661.

Digital Medicine Society (DiME) (2021) *DiMe's Library of Digital Endpoints*. Available at: *https://www.dimesociety.org/communication-education/library-of-digital-endpoints/*.

Eren-Oruklu, M. *et al*. (2009) 'Estimation of Future Glucose Concentrations with Subject-Specific Recursive Linear Models', *Diabetes Technology & Therapeutics,* 11(4), pp. 243–253. doi: 10.1089/dia.2008.0065.

Facchinetti, A., Sparacino, G. and Cobelli, C. (2011) 'Online Denoising Method to Handle Intraindividual Variability of Signal-to-Noise Ratio in Continuous Glucose Monitoring', *IEEE Transactions on Biomedical Engineering,* 58(9), pp. 2664–2671. doi: 10.1109/TBME.2011.2161083.

Gaynanova, I. and Li, G. (2017) 'Structural Learning and Integrative Decomposition of Multi-View Data'. Available at: *http://arxiv.org/abs/1707.06573*.

Georgiev, A. A., Ramsay, J. and Silverman, B. (1998) 'Functional Data Analysis', *Technometrics,* 40(3), p. 260. doi: 10.2307/1271190.

Goldsmith, J. *et al*. (2016) 'New Insights into Activity Patterns in Children, Found Using Functional Data Analyses', *Medicine & Science in Sports & Exercise,* 48(9), pp. 1723–1729. doi: 10.1249/MSS.0000000000000968.

Karas, M. *et al*. (2019) 'Accelerometry Data in Health Research: Challenges and Opportunities', *Statistics in Biosciences,* 11(2), pp. 210–237. doi: 10.1007/s12561-018-9227-2.

Li, G. and Gaynanova, I. (2017) 'A General Framework for Association Analysis of Heterogeneous Data', *arXiv preprint arXiv:1707.06485*.

Lock, E. F. *et al*. (2013) 'Joint and individual variation explained (JIVE) for integrated analysis of multiple data types', *The Annals of Applied Statistics*, 7(1), pp. 523–542. doi: 10.1214/12-AOAS597.

Mahalingaiah, S. *et al*. (2021) 'Design and methods of the Apple Women's Health Study: a digital longitudinal cohort study', *American Journal of Obstetrics and Gynecology. d*oi: 10.1016/j.ajog.2021.09.041.

Marler, M. R. *et al*. (2006) 'The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes', *Statistics in Medicine,* 25(22), pp. 3893–3904. doi: 10.1002/sim.2466.

Merikangas, K. R. *et al.* (2019) 'Real-time Mobile Monitoring of the Dynamic Associations Among Motor Activity, Energy, Mood, and Sleep in Adults With Bipolar Disorder', *JAMA Psychiatry,* 76(2), p. 190. doi: 10.1001/jamapsychiatry.2018.3546.

Nweke, H. F. *et al.* (2018) 'Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges', *Expert Systems with Applications,* 105, pp. 233–261. doi: 10.1016/j.eswa.2018.03.056.

Prendin, F. *et al*. (2021) 'Forecasting of Glucose Levels and Hypoglycemic Events: Head-to-Head Comparison of Linear and Nonlinear Data-Driven Algorithms Based on Continuous Glucose Monitoring Data Only', *Sensors,* 21(5), p. 1647. doi: 10.3390/s21051647.

Quer, G. *et al.* (2021) 'Wearable sensor data and self-reported symptoms for COVID-19 detection', *Nature Medicine,* 27(1), pp. 73–77. doi: 10.1038/s41591-020-1123-x.

Rabby, M. F. et al. (2021) 'Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction', *BMC Medical Informatics and Decision Making,* 21(1), p. 101. doi: 10.1186/s12911-021-01462-5.

Spira, A. P. et al. (2021) 'Brain Amyloid Burden, Sleep, and 24-Hour Rest/Activity Rhythms: Screening Findings from A4 and LEARN', *SLEEP Advances*. doi: 10.1093/sleepadvances/zpab015.

Uddin, M. Z. and Soylu, A. (2021) 'Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning', *Scientific Reports,* 11(1), p. 16455. doi: 10.1038/s41598-021-95947-y.

US FDA (2021) *Digital Health Technologies for Remote Data Acquisition in Clinical Investigations.*

Varma, V. R. et al. (2017) 'Total volume of physical activity: TAC, TLAC or TAC($λ$).', *Preventive medicine.* doi: 10.1016/j.ypmed.2017.10.028.

Yuda, E. et al. (2017) 'Sleep Stage Classification by a Combination of Actigraphic and Heart Rate Signals', *Journal of Low Power Electronics and Applications,* 7(4), p. 28. doi: 10.3390/jlpea7040028. ■

STATISTICAL CHALLENGES OF MODELLING OF MOBILE DIGITAL HEALTH DATA

Vadim Zipunnikov (Johns Hopkins), Debangan Dey (Johns Hopkins), Kathleen Merikangas (NIMH), Andrew Leroux (U Colorado, Denver)

INTRODUCTION Multi-sensor wearable devices are now used to perform real-time tracking of physical activity, sleep, heart rate, blood glucose, ambient and core body temperature, light and noise exposure and many more to come. This passive real-time tracking of human physiology and ambient environmental exposure is increasingly collected concurrently with data concerning individuals' behavior and mood state using self reported surveys administered through smartphone apps, typically multiple times per day, in a process referred to as Ecological Momentary Assessment (EMA), daily electronic diaries, experience sampling, or intensive longitudinal data. Historically, these surveys, hereafter referred to as EMA have collected information on varied aspects of human behaviours, including diet choices, self-perceived levels of energy, emotional states, physical activity, and quality of sleep (Csikszentmihalyi, 2011). EMA has been used to assess incident health and behavioral events, such as episodes of alcohol and substance use and abuse, headache or cardiac events, compliance and outcome of interventions (Bolger and Laurenceau, 2013). As an illustrative example, Figure 1 shows a multimodal sensor assessment in NIMH Family Study (Merikangas et al, 2019) that includes two sensors (actigraphy and light) coupled with multiple daily EMA assessments on participants' quality of sleep, mood, and energy. This combined assessment using passive multi-sensor monitoring and active EMA reporting creates mobile digital health data that provides tremendous opportunities to better understand human health and behavior, and to better inform prevention and intervention efforts. There is, however, a large gap between the complexity of mobile digital health data and statistical methodology designed for fully leveraging the potential for these data to provide insights into temporal dynamics between these various data streams and health outcomes. Below, we outline key statistical challenges of jointly modelling actigraphy and EMA.

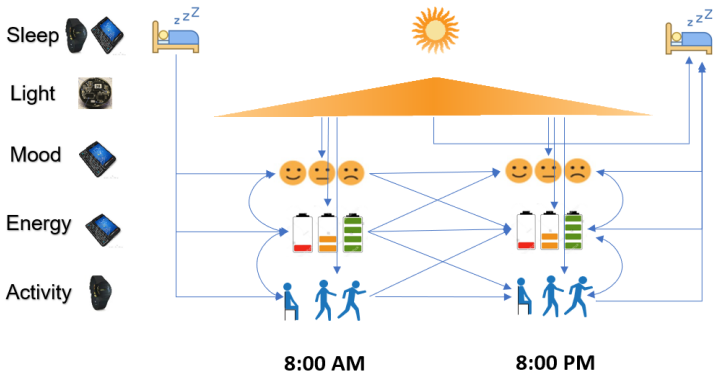


Figure 1. A schematic diagram of within-day inter-relationships between sleep (self-reported and actigraphy-estimated), light, self-reported mood and energy, and actigraphy-estimated motor activity.

ACTIGRAPHY The most commonly employed methods for analyzing actigraphy data fall well short of the sophistication of the data, specifically in the form of massive data reduction and loss of temporal information. These methods often rely on simple multi-day summaries which correspond to one of three domains measurable by wearable accelerometers: sleep, circadian rhythmicity, and physical activity. Moreover, studies generally only focus on features related to one of these three domains separately, when each has been shown to be a contributor to health outcomes. Recently, there has been increasing interest in methods which capture infor-

mation about temporal (or functional), distributional, and time-series aspects of actigraphy data. For example, functional data analysis (FDA) methods can be used to account for the temporal aspect of actigraphy data. FDA treats 24-hour actigraphy profiles as functional observations (Wang et al, 2016) and studies associations of health outcomes with both the timing and volume of actigraphic motor activity using, for example, scalar-on-function or function-on-scalar regression models (Goldsmith et al, 2016; Morris, 2015). FDA methods are fully data-driven without requiring parametric assumptions about the functional form of diurnal actigraphy profiles. The distributional aspect in actigraphy data can be captured via Distributional Data Analysis (DDA), a family of novel statistical methods that employ all subject-specific observations via either subject-specific probability distribution functions or quantile functions (Ghosal et al, 2021) that can then be used as predictors or outcomes (Petersen et al, 2022). For example, subject-specific activity bout durations during monitoring period are typically summarized as the mean activity bout duration and associated with outcomes of interest using classical regression. In contrast, DDA can summarize all bout durations through a subject-specific quantile function, where each quantile is associated with outcomes in a principled fashion. The time-series aspect can be captured using various time-series techniques. For example, Hidden Markov models with subject-specific time-varying parameters can be used to capture transition in and out of bouts of low/high activity (Lafan et al, 2010; Di et al, 2017). Novel integrative techniques such as Joint and Individual Variation Explained (JIVE) can simultaneously fuse multiple scalar features from each of the three domains of physical activity, sleep, and circadian rhythmicity (Di et al, 2019) or aggregate multiple aspects of actigraphy data including temporal, distributional, and time-series via joint and individual latent components (Varma et al, 2021; Ghosal et al, 2021). Differences in preferred timing of sleep and chronotypes can be accounted for using co-registration techniques, ensuring the amplitude and phase variations are properly separated (McDonnell et al, 2021). In addition to diurnal modelling, FDA approaches can account for weekly and seasonal trends which often present in actigraphy data but are commonly ignored in analyses (Wrobel et al, 2021). For cross-study analyses, additional challenges include differences in collection protocols, placements of devices, duration of wear time, even more so in combining data from studies that have employed different devices (Karas et al, 2022).

EMA measures typically exhibits a substantial between and within subject variability. To account for correlation in intensive longitudinal EMA measures, they are usually analyzed with multi-level methods (Bolger and Laurenceau, 2013), such as mixed effects models that take into account participant heterogeneity by allowing subject-specific random effects. Many key statistical challenges in EMA are similar to those in actigraphy and include: 1) informative and non-informative missingness that are frequently confounded by the contextual change in mental/physical status of individuals; 2) differences in subjective interpretation of scales; 3) cross-dependence across multiple EMA measures such as mood and energy; 4) presence of diurnal, weekly, and seasonal trends; 5) small sample sizes of many studies due to the intensity of assessments (Walls et al, 2006). In addition, many asssesments have different time scales (once per day, four times per day, or 24-hour), different timing of assessments (e.g. four fixed times per day vs four subject-specific times per day) and many EMA measures are collected using different measurement scales (binary, nominal, ordinal, truncated, or continuous) (Dunster et al, 2021).

JOINT MODELLING OF ACTIGRAPHY AND EMA inherits all modality-specific challenges and is further complicated by significant cross-dependences across multiple modalities. Recent developments in joint modelling of Actigraphy and EMA data primarily focused on non-structured machine learning methods (Kim et al, 2019) and more structured approaches such as dynamical structural equation models (Asparouhov et al, 2018; Merikangas et al, 2019). Identifying a structural change (Aminikhanghahi and Cook, 2017) in combined actigraphy and EMA data requires simultaneous exploration and monitoring multiple aspects present in these two modalities - distributional, temporal, time-series in actigraphy and mean, variability, stability, autocorrelation, fragmentation in EMA (Johns et al, 2019).

HARMONIZATION The process of transforming mobile digital health data into knowledge is impossible without active intellectual participation of statisticians and subject matter experts in major multidisciplinary efforts that focus on conceptualization, measurement, analysis and treatment of multiple physiological, behavioral, and health components. One example of such an effort is the "mobile Motor Activity Research Consortium for Health" (mMARCH), a collaborative network of studies of clinical and community samples that employ common mobile, clinical, and biological measures to examine the generalizability and clinical significance of findings across involved studies (Gideon et al, 2021). mMARCH currently includes sites both in the US and across the world. The main scientific goals of mMARCH sites are broad, ranging from developing an understanding of the inter-relationships of physical activity, sleep, and mood to the dynamic interplay between sleep, stress, alcohol and substance use, and other health behaviors in children, adolescents, and adults. These goals can only be accomplished through the combined, large-scale effort, involvement, and leadership of statisticians whose expertise in the many relevant areas of statistics will allow for progress on the fronts of standardization of data collection protocols, sampling design, and application and development of novel analytical methods.

CONCLUSION The development of novel statistical methods should be accompanied by substantial collaborative efforts with behavioral scientists and clinicians to translate the powerful potential of mobile digital health data into actionable health information.

Contact email: *vadim.zipunnikov@gmail.com*

References

Aminikhanghahi, S. and Cook, D.J., 2017. A survey of methods for time series change point detection. *Knowledge and information systems,* 51(2), pp.339-367.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W. and Pollak, C.P., 2003. The role of actigraphy in the study of sleep and circadian rhythms. Sleep, 26(3), pp.342-392.

Asparouhov, T., Hamaker, E.L. and Muthén, B., 2018. Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), pp.359-388.

Bolger, N. and Laurenceau, J.P., 2013. *Intensive longitudinal methods: An introduction to diary and experience sampling research.* Guilford press.

Csikszentmihalyi, M. (2011). *Handbook of research methods for studying daily life.* Guilford Press.

Di, J., Leroux, A., Urbanek, J., Varadhan, R., Spira, A.P., Schrack, J. and Zipunnikov, V., 2017. Patterns of sedentary and active time accumulation are associated with mortality in US adults: The NHANES study. *bioRxiv*, p.182337.

Di, J., Spira, A., Bai, J., Urbanek, J., Leroux, A., Wu, M., Resnick, S., Simonsick, E., Ferrucci, L., Schrack, J. and Zipunnikov, V., 2019. Joint and individual representation of domains of physical activity, sleep, and circadian rhythmicity. *Statistics in biosciences,* 11(2), pp.371-402.

Dunster, G.P., Swendsen, J. and Merikangas, K.R., 2021. Real-time mobile monitoring of bipolar disorder: a review of evidence and future directions. *Neuropsychopharmacology,* 46(1), pp.197-208.

Johns, J. T., Di, J., Merikangas, K., Cui, L., Swendsen, J., & Zipunnikov, V. (2019). Fragmentation as a novel measure of stability in normalized trajectories of mood and attention measured by ecological momentary assessment. *Psychological Assessment,* 31(3), 329.

Ghosal, R., Varma, V. R., Volfson, D., Hillel, I., Urbanek, J., Hausdorff, J. M., ... & Zipunnikov, V. Distributional data analysis via quantile functions and its application to modeling digital biomarkers of gait in Alzheimer's Disease. Biostatistics.

Goldsmith, J., Liu, X., Jacobson, J. and Rundle, A., 2016. New insights into activity patterns in children, found using functional data analyses. Medicine and science in sports and exercise, 48(9), p.1723.

Karas, M., Muschelli, J., Leroux, A., Urbanek, J.K., Wanigatunga, A.A., Bai, J., Crainiceanu, C.M. and Schrack, J.A., 2022. Comparison of accelerometry-based measures of physical activity. medRxiv.

Kim, H., Lee, S., Lee, S., Hong, S., Kang, H. and Kim, N., 2019. Depression prediction by using ecological momentary assessment, actiwatch data, and machine learning: observational study on older adults living alone. JMIR mHealth and uHealth, 7(10), p.e14149.

Laffan, A., Caffo, B., Swihart, B. J., & Punjabi, N. M. (2010). Utility of sleep stage transitions in assessing sleep continuity. Sleep, 33(12), 1681-1686.

McDonnell, E.I., Zipunnikov, V., Schrack, J.A., Goldsmith, J. and Wrobel, J., 2021. Registration of 24-hour accelerometric rest-activity profiles and its application to human chronotypes. Biological Rhythm Research, pp.1-21.

Merikangas, K. R., Swendsen, J., Hickie, I. B., Cui, L., Shou, H., Merikangas, A. K., ... & Zipunnikov, V. (2019). Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. JAMA psychiatry, 76(2), 190-198.

Morris JS. Functional regression. Annual Review of Statistics and Its Application. 2015 Apr 10;2:321-59.

Petersen, A., Zhang, C. and Kokoszka, P., 2022. Modeling probability density functions as data objects. Econometrics and Statistics, 21, pp.159-178.

Varma, V.R., Ghosal, R., Hillel, I., Volfson, D., Weiss, J., Urbanek, J., Hausdorff, J.M., Zipunnikov, V. and Watts, A., 2021. Continuous gait monitoring discriminates community□dwelling mild Alzheimer's disease from cognitively normal controls. Alzheimer's & Dementia: Translational Research & Clinical Interventions, 7(1), p.e12131.

Walls, T.A. and Schafer, J.L. eds., 2006. Models for intensive longitudinal data. Oxford University Press.

Wang, J.L., Chiou, J.M. and Müller, H.G., 2016. Functional data analysis. Annual Review of Statistics and Its Application, 3, pp.257-295.

Wrobel, J., Muschelli, J. and Leroux, A., 2021. Diurnal physical activity patterns across ages in a large UK based cohort: the UK biobank study. Sensors, 21(4), p.1545. ■

ACCELERATING MRNA PROCESS DEVELOPMENT WITH DIGITAL TWINS

Elliott Schmitt (Moderna), Gang Wang (Moderna), Julia O'Neill (Moderna)

I. mRNA manufacturing platform enables rapid response to urgent demands

In 2020, the sudden onset of the global SARS-CoV-2 pandemic crisis required a fast response to avert tragic losses. Vaccine development timelines, which can take 10-15 years (Offit, P. MD, 2020) from discovery to approval, were accelerated. The compressed timelines required many activities, such as clinical trials, process development, and health authority reviews to be completed in parallel, instead of in the traditional sequential order (Figure 1). The accelerated product timeline enabled emergency use authorization for COVID vaccines from both Pfizer/BioNTech and Moderna in less than a year. Both the Pfizer/BioNTech and Moderna vaccines were based on mRNA technology and successfully demonstrated the speed, flexibility, and scalability of the platform for vaccine manufacturing.

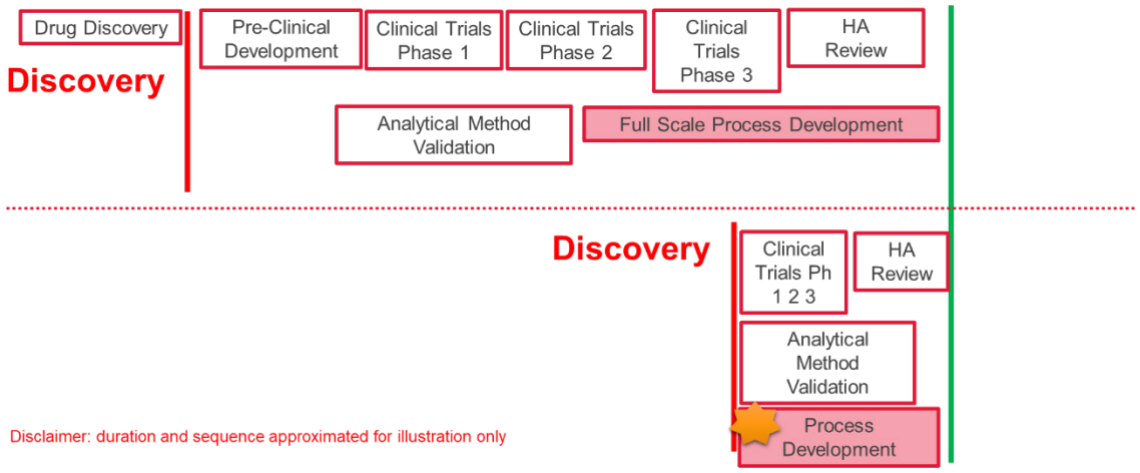


Figure 1 (Top) Traditional approach to vaccine development, (Bottom) Accelerated approach to vaccine development

The commercialization of the mRNA-based vaccine platform approach for COVID has opened the door to future mRNA-based medicines. A platform process is an approach of developing a production strategy for a new drug based on established manufacturing processes (Q11 Development and Manufacture of Drug Substances, 2012). Once the platform is proven through the acceptance of an initial drug product (e.g. Moderna mRNA-1273, Spikevax), a template for future drug products is created. Each new program follows an iterative process, building upon the knowledge from the previous programs, incorporating new data and innovations into the next-generation platform. This platform-based approach enables multiple product opportunities, a higher probability of technical success, accelerated research and development timelines and greater capital efficiency with each new product.

The principles of Quality by Design (QbD) are complementary to platform processes. A QbD approach to development emphasizes understanding and control of process parameters and product quality attributes. This can be achieved through using statistical design of experiments (DoE), process modeling, risk assessment, and process analytical technology (PAT) to characterize the relationship between Critical Process Parameters (CPP) and Critical Quality Attributes (CQA). Knowledge of these relationships is used to determine the design space and define a control strategy to assure product

quality. The time, resource, and financial investment costs to establish a QbD control strategy can be significant. Process development can take 3 – 5 years and cost anywhere from $40 – 100 million per product (Farid et al., 2020). With a platform process, prior knowledge can be used to support design decisions and reduce some of the development burden for next-generation products. This can free up resources to take advantage of more sophisticated tools to increase process understanding and improve product quality.

II. Accelerated process development through Digital Twins

Digital twins have been presented as a promising approach to accelerate process development (Cardillo et al., 2021; Schmidt et al., 2021; Zobel-Roos et al., 2019). A digital twin can be defined as a virtual representation that serves as a real-time digital counterpart of a physical process. The main components of a digital twin include a process model, digital infrastructure (databases), and real-time sensors (Figure 2). The process can be defined by the system dynamics, process constraints (or controls) and target states. Using Digital twins for in silico process development could save 50 – 75% in process development time (Stosch et al., 2021). If a digital twin for an mRNA-based platform process could reduce process development time by 50%, then all BLA required activities could be completed in less than 6 months. This substantial benefit of acceleration motivates future process development work for mRNA-based vaccines and therapeutics to rely on using digital twins to drive *in silico* process design, optimization, and validation.

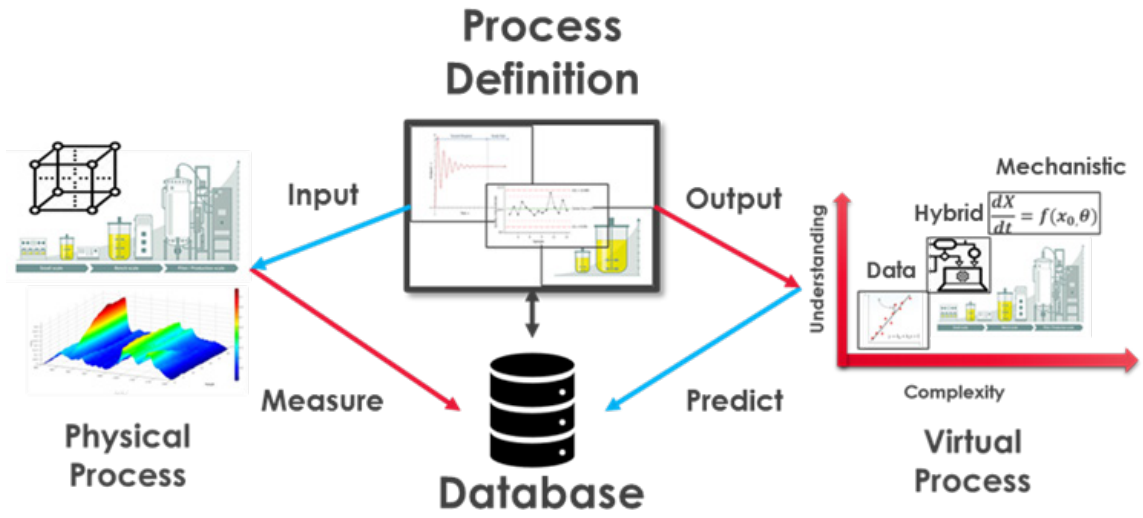


Figure 2 The components of a Digital Twin of a manufacturing process

III. Case study: AEX model development and application

In what follows, an AEX (anion-exchange) chromatography process is used as a case study to describe how a Digital Twin has been implemented and applied to perform *in silico* design of experiments (DoE) for process robustness study at Moderna. An overview of the general workflow is illustrated in Figure 3. First, the structural form of the mechanistic model is defined for the AEX process based on domain expertise and previous experience. Then, deliberately designed experiments are used to calibrate, or determine the parameters of, the mechanistic model. After the AEX model has been

calibrated, it is ready to be used for *in silico* DoE, process design and optimization, operational decisions, and other applications.

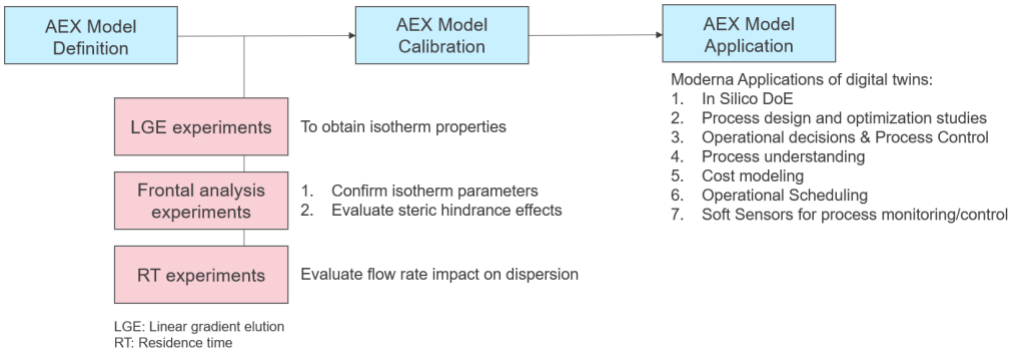


Figure 3 Workflow of digital twin creation combining in silico and lab experiments

For the AEX case study, the classic equilibrium dispersive model (Figure 4) is used to describe the mass transfer in the chromatography column (Guélat et al., 2016; Khalaf et al., 2016; Ng et al., 2012). In this model, concentrations in both mobile phase C and stationary phase $\langle\hat{q}\rangle$ are considered functions of only axial position (x) and time (t). Term $\langle\hat{q}\rangle$ represents the average concentration within a resin particle including the porous liquid: $\langle\hat{q}\rangle = \epsilon_{part}\bar{C}_p + (1-\epsilon_{part})\bar{q}$. In this model, column porosity ϵ_{col} and axial dispersity D_{ax} are both parameters indicating column packing quality. u_{int} is the interstitial linear velocity, which is determined by operating conditions. In most cases, only numerical solution of this partial differential equation (PDE) system is necessary for industrial applications. A large variety of commercial and free PDE solvers are available. For this case study, GoSilico™ Chromatography Modeling Software, offered by Cytiva, has been used as the tool for model calibration and application.

$$\frac{\partial C}{\partial t} + u_{int}\frac{\partial C}{\partial x} = D_{ax}\frac{\partial^2 C}{\partial x^2} - \frac{(1-\epsilon_{col})}{\epsilon_{col}}\frac{d\langle\hat{q}\rangle}{dt}$$

Figure 4 Mass transfer equation (equilibrium dispersive) of column chromatography

Linear or nonlinear optimization is used for model calibration, where parameters are estimated to minimize the residuals between the model predictions and experimental observations. Scores like mean squared error or mean absolute error are commonly used to locate the optimum (minimal) residuals. It is critical to select experimental run settings to support calibration for the specific mechanistic model form to be calibrated. In contrast to empirical statistical models based on linearized Taylor expansion model terms, mechanistic models are often highly nonlinear, and the residual sensitivity to model parameters is not always trivial. Therefore, classical DoE design matrices, such as full factorial or central composite designs, may not support optimal parameter fitting efficiency and estimation confidence. Model-based DoE approaches are commonly used to fit mechanistic models (Shahmohammadi & McAuley, 2019). For chromatography models, experimental designs relying on linear gradient elution (LGE), isocratic elution, and frontal analysis experiments are commonly used to calibrate IEX (ion-exchange) models (Carta et al., 2020, 2005). In this approach, linear gradient elution experiments with different gradients are performed first to determine the isotherm properties. Then, frontal analysis is used to validate the isotherm properties and evaluate parameters describing steric hindrance effects and capacities. Finally, axial dispersion dependency on the flow rate is evaluated by running either gradient elution or isocratic elution experiments with different residence times. Global optimization

techniques are applied to fit the model parameters. With this dedicated setup of experiments, satisfactory agreement is reached and only reached when 1) the mechanistic model captures all physics in the system, and 2) optimal model parameters are found (Figure 5).

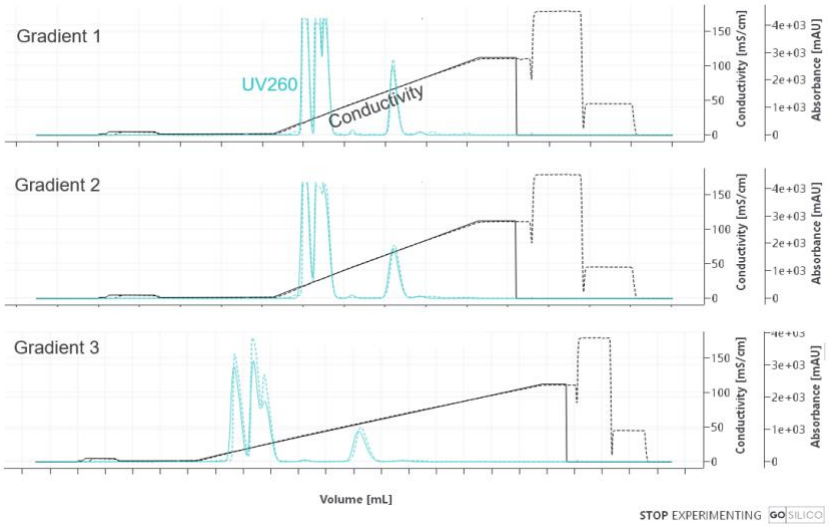


Figure 5 Model calibration using deliberately designed experiments and agreement is found between model predictions and experiments (screenshot from GoSilico™ Chromatography Modeling Software)

The mechanistic model can be applied for *in silico* predictions when domain experts endorse the model performance. *in silico* experiments can explore design space where lab experiments are challenging to perform due to limitations such as material availability and time. As an example, we used the AEX chromatography model to evaluate and identify critical material attributes (CMAs) from upstream unit operations. In lab experiments, it is nearly impossible to generate abundant samples with combinations of concentrations of multiple components, and the time to perform such experiments in this high-dimensional space is prohibitive. Using the Digital Twin we developed, we generated 1000 parameter sets in a high-dimensional design space using a Latin hypercube approach (Figure 6) and ran *in silico* simulations. Critical material attributes were identified and the simulation results will give us insight on upstream specification setting.

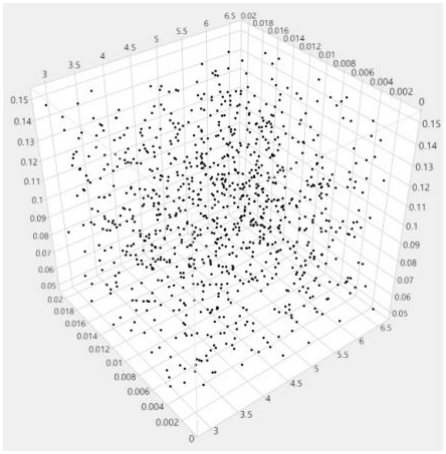


Figure 6 Latin hypercube / Monte Carlo generated design space for in silico simulations (scatter plot generated by JMP)

IV. Model-based design and filing

Most pharmaceutical development and regulatory filings applying Quality by Design (QbD) principles are based on classical design of experiments and empirical statistical models described in the ICH guidance document Q8(R2) (ICH Q8(R2), 2009). Predictions are interpretable and robust within the design space, but uncertainty increases when the process parameters extend beyond the design space or factors that are not accounted for during process characterization turn out to have impact on the process. The original experiments can be augmented empirically to address these situations. However, Digital Twins and mechanistic models provide more reliable predictions when extending the design space and incorporating factors not considered during the original characterization. Their strength relies on the basic principle that extrapolation of mechanistic models beyond the original design space is more reliable than extrapolation beyond the demonstrated limits of empirical models.

Digital Twins increase the confidence in the statistical models built upon process characterization studies by explaining the mechanism of the significant process parameters. Developing capability of *in silico* prediction helps to target experiments on parameter spaces where there is greater information potential, such as higher risk of quality attributes failing specifications or greater likelihood of global rather than local optima. Instead of relying solely on domain expertise in defining the range of design space, rational determination of the space can be achieved by employing mechanistic models. The combination of Digital Twins and statistical model adds even greater value to platform processes, where mRNA sequence dependent information could be incorporated into the mechanistic model and targeted experiments could be performed to assess risks at certain region of the design space and ensure the QbD principles.

Acknowledgement

The authors would like to thank Matthew Burak and Haripriya Sharma for performing the experiments and Chris Ladd for advice and guidance on chromatography modeling.

Conflict of interest

All authors are employees of and shareholders in Moderna, Inc. The authors declare no other competing interests.

References

Cardillo, A. G., Castellanos, M. M., Desailly, B., Dessoy, S., Mariti, M., Portela, R. M. C., Scutella, B., Stosch, M. von, Tomba, E., & Varsakelis, C. (2021). Towards in silico Process Modeling for Vaccines. *Trends in Biotechnology*, *39*(11), 1120–1130. https://doi.org/10.1016/j.tibtech.2021.02.004

Carta, G. & Jungbauer, A. (2020). *Protein chromatography: process development and scale-up*.

Carta, G., Ubiera, A. R., & Pabst, T. M. (2005). Protein Mass Transfer Kinetics in Ion Exchange Media: Measurements and Interpretations. *Chemical Engineering & Technology*, *28*(11), 1252–1264. https://doi.org/10.1002/ceat.200500122

Farid, S. S., Baron, M., Stamatis, C., Nie, W., & Coffman, J. (2020). Benchmarking biopharmaceutical process development and manufacturing cost contributions to R&D. *MAbs*, *12*(1), 1754999. https://doi.org/10.1080/19420862.2020.1754999

Guélat, B., Khalaf, R., Lattuada, M., Costioli, M., & Morbidelli, M. (2016). Protein adsorption on ion exchange resins and monoclonal antibody charge variant modulation. *Journal of Chromatography A*, *1447*, 82–91. https://doi.org/10.1016/j.chroma.2016.04.018

ICH. (2009). *Q8(R2) Pharmaceutical Development*.

ICH. (2012). *Q11 Development and Manufacture of Drug Substances*

Khalaf, R., Heymann, J., LeSaout, X., Monard, F., Costioli, M., & Morbidelli, M. (2016). Model-based high-throughput design of ion exchange protein chromatography. *Journal of Chromatography A*, *1459*, 67–77. https://doi.org/10.1016/j.chroma.2016.06.076

Ng, C. K. S., Osuna-Sanchez, H., Valéry, E., Sørensen, E., & Bracewell, D. G. (2012). Design of high productivity antibody capture by protein A chromatography using an integrated experimental and modeling approach. *Journal of Chromatography B*, *899*, 116–126. https://doi.org/10.1016/j.jchromb.2012.05.010

Offit, P. (2020). *Vaccine Development, Testing, and Regulation*. Retrieved March 8, 2022, from https://www.historyofvaccines.org/content/articles/vaccine-development-testing-and-regulation

Schmidt, A., Helgers, H., Vetter, F. L., Juckers, A., & Strube, J. (2021). Digital Twin of mRNA-Based SARS-COVID-19 Vaccine Manufacturing towards Autonomous Operation for Improvements in Speed, Scale, Robustness, Flexibility and Real-Time Release Testing. *Processes*, *9*(5), 748. https://doi.org/10.3390/pr9050748

Shahmohammadi, A., & McAuley, K. B. (2019). Sequential model-based A- and V-optimal design of experiments for building fundamental models of pharmaceutical production processes. *Computers & Chemical Engineering*, *129*, 106504. https://doi.org/10.1016/j.compchemeng.2019.06.029

Stosch, M. von, Portela, R. M., & Varsakelis, C. (2021). A roadmap to AI-driven in silico process development: bioprocessing 4.0 in practice. *Current Opinion in Chemical Engineering*, *33*, 100692. https://doi.org/10.1016/j.coche.2021.100692

Zobel-Roos, S., Schmidt, A., Mestmäcker, F., Mouellef, M., Huter, M., Uhlenbrock, L., Kornecki, M., Lohmann, L., Ditz, R., & Strube, J. (2019). Accelerating Biologics Manufacturing by Modeling or: Is Approval under the QbD and PAT Approaches Demanded by Authorities Acceptable Without a Digital-Twin? *Processes*, *7*(2), 94. https://doi.org/10.3390/pr7020094 ■

DIGITAL BIOMARKERS OF GLUCOSE CONTROL - REPRODUCIBILITY CHALLENGES AND OPPORTUNITIES

Irina Gaynanova (Texas A&M)

Introduction

The prevalence of diabetes, a chronic disease characterized by high blood glucose levels, continues to rise globally. According to the International Diabetes Federation, 537 million adults live with type 1 or type 2 diabetes, and this number is predicted to rise to 643 million by 2030 and 783 million by 2045[1]. Glycated hemoglobin A1c (HbA1*c*) has proven to be a reliable biomarker for diagnosing and monitoring the management of diabetes[2]. First, HbA1*c* is an indicator of long-term glucose control as it reflects the average blood glucose over the preceding 2-3 months. Second, high levels of HbA1*c* correlate with long-term diabetes complications, with HbA1*c* being an independent risk factor for coronary heart disease and stroke in subjects with or without diabetes. Despite this, HbA1*c* as a sole biomarker has severe limitations. Glucose levels are highly non-linear and non-stationary, even for healthy subjects. They are strongly affected by various environmental factors, including diet, physical activity, stress, and sleep quality[3]. HbA1*c* is unable to capture this glycemic variability.

Continuous glucose monitors (CGM) are small wearable devices that automatically measure blood glucose levels at frequent time intervals, with some monitors taking measurements as often as every 5 min. Unlike HbA1*c*, CGM data provide a detailed quantification of the variation in blood glucose levels, thus playing an increasing role in clinical practice[4,5]. Figure 1 shows example glucose measurements for a subject with type 2 diabetes over 13 days obtained with Dexcom G4; these data are part of R package iglu[6]. Observe the highly non-stationary nature of the profile with multiple peaks (likely corresponding to meal intakes) and deviations from the in-range values. The increasing use of CGM devices coupled with their increasing measurement accuracy led to enormous interest in extracting CGM-based digital biomarkers of glucose control to replace or enhance the traditional HbA1*c* biomarker.

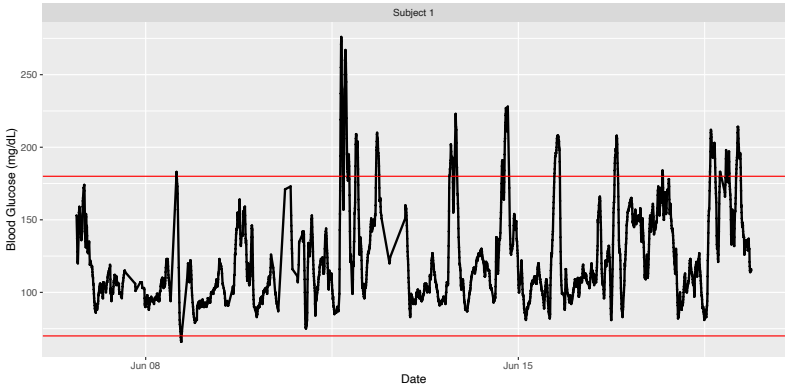


Figure 1: Glucose profile of a subject with Type 2 diabetes measured using Dexcom G4 CGM with 5 min frequency. The horizontal lines correspond to the in-range glucose thresholds of [70,180] mg/dL.

CGM-derived metrics as digital biomarkers of glucose control

Multiple CGM-derived digital biomarkers of glycemic control have been proposed[7,8]. These digital biomarkers can be viewed as summary statistics extracted from the complete CGM glucose profile aimed at measuring different glycemic aspects such as overall glucose levels (e.g., mean glucose), overall glucose variability (e.g., percent of glucose values in [70,180] mg/dL range, and coefficient of variation), local glucose variability (e.g., mean amplitude of glycemic excursions[9], and the standard deviation of glucose rate of change[10]), hypoglycemic (e.g., Hypo Index) and hyperglycemic (e.g., Hyper Index) risks. New CGM metrics continue to be proposed and developed, including composite metrics that account for multiple dimensions of glucose control (e.g., GRADE[11], and COGI[12]). The variety of available CGM metrics poses continuous challenges for researchers, clinicians, and patients alike.

On the one hand, in patient-centered care, it is crucial to use CGM biomarkers that have clear evidence of association with diabetes-associated risks and are easy to extract, interpret, and explain to the patient. Driven by these considerations, an international panel selected a small list of CGM metrics for clinical care, including mean glucose value, time in [70,180] mg/dL range, coefficient of variation[13]. On the other hand, in clinical

research, having accessibility to a broader range of metrics is of value due to the multi-faceted nature of glucose control (overall glucose levels, glucose variability, hypoglycemic risk, etc.). While metrics that provide a dynamic characterization of glucose trajectories (e.g., the mean amplitude of glycemic excursions MAGE, CONGA, and the standard deviation of rate of change) may be challenging to explain to the patient, they are relevant when considering the potential impact of conditions associated with acute temporal changes in pathophysiological mechanisms. However, there is a lack of consensus on the optimal metrics for assessing glucose variability[14] and no agreement on how to assign weights to different metrics of glucose control when determining disease progression[8].

We argue that there are two main challenges in utilizing CGM-derived metrics as digital biomarkers of glucose control. The first challenge is associated with the relative recency of CGMs, and as a result, limited CGM use. The newness of CGM technology translates into the lack of long-term prospective outcome studies for objective cross-comparison of various metrics and the lack of studies validating metrics' relationship to predicting long-term diabetes complications. This challenge is well-acknowledged within clinical literature[8] and can only be addressed with time as these studies take place. The second challenge, however, is more subtle and should be addressed now, but in our view, received much less attention - a challenge of reproducibility. Limited CGM software options, variation in algorithms used for metrics calculation, and lack of algorithms' validation on public datasets all lead to (a) disagreement in the values of the same CGM metrics across software platforms; (b) disregard of some metrics due to their inaccessibility rather than due to their potential clinical value. Statisticians and data scientists are well-positioned to help address these challenges, and here we review some efforts done by our group as well as outline opportunities for future work.

CGM reproducibility challenges and opportunities

1. Limited metrics' implementation in existing software solutions.

Ambulatory Glucose Profile (AGP) is the standardized report of CGM metrics offered through most CGM devices and their reporting software[13]. This report, however, is designed for individualized patient care and ease of interpretation rather than for large-scale research studies. AGP only includes a small subset of CGM metrics and, in particular, lacks metrics that provide a dynamic characterization of glucose trajectories (e.g., MAGE, and the standard deviation of rate of change). Furthermore, while some CGM metrics are easy to calculate without specialized software (e.g., average glucose value, and coefficient of variation), other metrics require non-trivial computations. For example, the average daily risk range (ADRR)[14] is a composite metric that requires both logarithmic and power transformations of glucose values, truncation of values based on specified hypo- and hyperglycemic cutoffs, and a weighted average. While the corresponding mathematical formula is explicit, we find it unlikely that non-quantitative researchers will utilize ADRR in the absence of readily available software. Since ADRR is not a part of standardized AGP, it is not routinely calculated; thus, its potential clinical utility (or lack of thereof) often remains unassessed in large-scale studies. The situation is even more complicated with time-specific measures of standard deviation such as SdB metric (between day – within time points standard deviation)[7]. Calculation of SdB requires putting glucose measurements on the same time grid from day to day (thus requiring interpolation or time adjustment), and then calculating various summary statistics across days and times. As with ADRR, we find it unlikely that clinicians will routinely calculate SdB without easily accessible software. In summary, there is a gap between CGM biomarkers proposed in the literature and CGM biomarkers routinely calculated in practice. There is a need for software tools that are free, easy to use for both researchers and clinicians, and up to date with the array of CGM biomarkers being proposed.

In recent years, multiple free software packages have been developed to address this need. EasyGV is a free CGM software in a macro-enabled Excel workbook[15] and thus is accessible to a broad range of clinicians and researchers. However, Excel is not a script-based programming language, and many underlying algorithms are black-box. It is thus less desirable for those users who want to create reproducible scripts for all data processing and metric calculation steps. It also has a limited array of available CGM metrics. A growing interest in open-source software for CGM data led to the development of multiple R packages within just the last two to three years[16, 17], including the iglu[6] package developed by our group. The main advantages of iglu over its

predecessors are (i) a significantly more exhaustive list of available CGM metrics; (ii) accompanying graphical user interface via Shiny app (*https://irinagain.shinyapps.io/shiny_iglu*) making it accessible to users with limited programming experience and (iii) community-based approach to maintenance and development of new features based on public GitHub repository to stay up to date with the most current CGM metrics. Given how frequently new CGM metrics appear in the literature, we believe open-source approaches coupled with community efforts are needed for any CGM software solution to stay relevant. Having easy access to well-documented and tested software for calculating continuously proposed CGM metrics is crucial for determining their clinical importance in follow-up studies, reproducibility, and follow-up analyses of interventions' effects on glucose control.

2. Proprietary algorithms and black-box data processing.

The algorithms and data processing techniques used in various reporting systems provided by CGM manufacturers are proprietary, making the comparisons difficult, and often leading to disagreements across software systems. Lack of agreement is especially challenging for metrics whose calculation is non-trivial due to (a) required underlying data-processing and (b) algorithmic differences in translation of clinical metric definition.

By default, CGMs take measurements on an equidistant time grid. However, missing values are common due to device placement issues, device replacement from one to another, and underlying glucose values outside the CGM measurement range (too low or too high). In the presence of such missingness, each software implementation must either ignore these missing values or impute them, with the latter requiring an additional decision on the specific imputation scheme. As many existing software and algorithms are proprietary, it is unclear what choice is being made. Even with open-source software, the explicit choice is not always adequately described in the documentation, leading to the lack of reproducibility across the implementations.

But even in the absence of missing data, significant variations in underlying algorithms could exist. For example, the mean amplitude of glycemic excursions (MAGE) is a commonly used measure of glucose variability. Based on the original definition[9], MAGE is the arithmetic mean of the amplitude (height) of glucose excursions greater than the standard deviation of the glucose values. Visually, an "excursion" is a substantial peak in glucose values, typically due to a meal intake. Translation of MAGE definition into an automated algorithm requires quantification of this visual "excursion". In our experience, the variability in this quantification leads to differences in calculated MAGE values. These differences could be substantial across software platforms and even used as a rationale for abandoning MAGE as a metric altogether[18]. But to us, this rationale seems faulty. The issue is not MAGE's clinical utility but rather a choice of the algorithm for automatically identifying excursions that will maximize this utility. As statisticians, we are accustomed to doing simulation studies to compare various algorithms on either synthetic data or on actual public data so that comparisons can be replicated and verified by others. We argue that the same approach is desperately needed to evaluate and compare CGM software and associated algorithms. In the context of MAGE, we developed an open-source algorithm for excursion identification[19] with implementation being freely available through R package iglu. As a standard of comparison, we used visual identification of excursions and manual calculations on the public rather than private CGM data. While manual calculations are undoubtedly error-prone, the public availability of all software, underlying algorithms, and the utilized CGM data ensures that the results can be publicly disproved or validated by the community. We believe such comparison efforts are desperately needed for many other CGM metrics and CGM software and can benefit from the more active involvement of quantitative scientists. It is encouraging to see some recent progress in this area[20, 21,] and we hope that these efforts continue.

3. Lack of easily accessible public CGM data for validation of metric calculation and development.

Most of the studies proposing new CGM metrics and software rely on CGM data that is not publicly available. Lack of data availability makes it difficult to validate the correctness of metrics' calculations, hindering the development of better algorithms and software. We believe that the area of CGM digital biomarkers research has some lessons to learn from the machine learning community. The availability of a central public repository of datasets to test the performance of various algorithms, widely known as UCI machine learning repository [22], has mainly been credited for the exponential growth in machine learning algorithms development. The standard in the field is not to display the results on only private data that no one has access

to, but to display the results on public data that everyone can access, and thus use as a benchmark. Public availability of deidentified CGM datasets and ease of access will allow evaluation of the agreements and differences across CGM software and the effect of those disagreements on the relationship between calculated metrics and clinical outcomes of interest. We believe that such public availability will help to naturally filter both algorithms and CGM metrics, as well as generate more confidence in the community in the correctness of underlying calculations. Our group has made some effort towards this goal by assembling links to public CGM data with corresponding processing scripts in R and python hosted on GitHub[23], with contributions open to the community. We believe having easy access to public deidentified CGM datasets will spearhead the validation of CGM biomarkers and ensure their reproducibility.

4. Lack of large-scale data-driven studies to assess the relationship between the metrics.

While using multiple indices from CGM data allows considering different characteristics of glucose control, many of the proposed digital biomarkers appear to be redundant. For example, glucose management indicator (GMI)[24] is a linear function of mean glucose. While it provides utility in terms of its values being directly comparable to HbA1*c* values, it does not give additional information on glucose control beyond the information provided by the mean glucose. Traditionally, an expert opinion is needed to differentiate redundant metrics and identify which aspects of glucose control (overall levels, global variability, local variability, etc.) various metrics represent. However, such expert opinion is increasingly challenging to obtain as new metrics continue to be developed. It's easy to be overwhelmed by the sheer amount of information across all metrics, much of which is overlapping. There is a need to evaluate metrics similarity and redundancy in a data-driven way to complement existing clinical expertise.

To illustrate the idea, in Figure 2, we reproduce Figure 6 from Broll et al.[6] showing hierarchical clustering of 40+ CGM metrics based on public data of five subjects with Type 2 diabetes available in iglu package[6]. The cluster tree for metrics is cut at 6 groups, which we interpret as follows (from top to bottom):

(1) In range metrics;
(2) Hypoglycemia metrics;
(3) Hyperglycemia metrics;
(4) A mixture of variability and hyperglycemia metrics;
(5) CVsd (standard deviation of CV, coefficient of variation, across days);
(6) Glucose variability metrics.

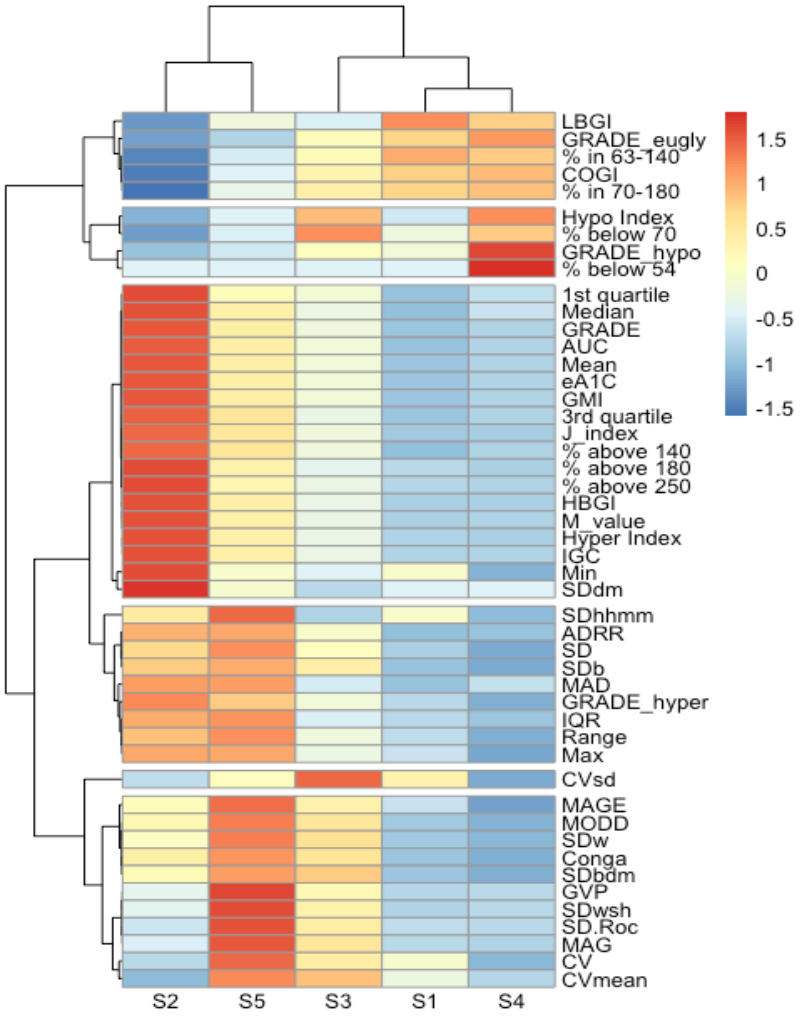


Figure 2. Heatmap of CGM metrics calculated using R package iglu for five subjects with Type 2 diabetes. Hierarchical clustering is performed on centered and scaled metric values using distance correlation with complete linkage. The cluster tree for metrics is cut such that it results in six groups.

Such visual representation provides a data-driven way to assess metrics' similarity and can be used as a guide for metrics selection. Metrics clustering across various populations (e.g., patients with Type 1 diabetes vs. patients with Type 2 diabetes vs. patients with gestational diabetes) may shed light on the condition-specific choice of digital biomarkers.

As an alternative approach to clustering, Fabris et al.[25, 26] investigated the use of sparse principal component analyses as a data-driven way to create new composite digital biomarkers from the existing metrics list. However, only a small subset of metrics has been considered, the studies were based on very small CGM datasets (<20 subjects each), and those datasets are not in public access.

There are great opportunities to use algorithms from statistics and machine learning (clustering, principal component analysis, factor analysis) to shed light on relationships between existing CGM metrics and quantify each metric's additional information. Significantly more large-scale studies are needed in this direction. Still, ultimately their success will depend on overcoming other reproducibility challenges first: comprehensive and publicly available implementation software, open-source algorithms, and public CGM datasets.

Conclusion

Continuous improvement in CGMs accuracy and their increased use provide exciting opportunities for improving glucose control of patients leaving with diabetes and improving our understanding of diabetes progression, risk factors, and treatment effects. However, multiple challenges exist in realizing these opportunities. Here we outlined four reproducibility challenges associated with the CGM-based digital biomarkers of glucose control: limited software implementations, proprietary algorithms, lack of easily accessible public CGM data, and lack of large-scale data-driven studies to assess the relationships across metrics. We believe that statisticians and data scientists are well-positioned to help address these challenges by developing algorithms, software, and reproducible workflows for CGM data processing and downstream analyses. Some examples of these efforts by our group are the free and open-source R package iglu[6] for the calculation of various CGM metrics, the accompanying Shiny app (*https://irina-again.shinyapps.io/shiny_iglu/*), and the curated GitHub repository with information on public CGM datasets[23]. Open-source implementation and curation of public CGM datasets for cross-comparison and benchmarking are all necessary to ensure validation by the research community at large. Significantly more efforts are needed in this direction, including initiatives that support collaborations between statisticians. data scientists, CGM manufacturers, and clinicians to ensure rigor, reproducibility, and clinical utility. As long-term prospective outcome studies collecting CGM data become a reality, reproducible analytical solutions are desperately needed to facilitate the use of CGM technology by investigators in both observational studies and randomized clinical trials.

References

1. International Diabetes Federation (2021). International Diabetes Federation: IDF Diabetes Atlas, 10th ed. Brussels. *www.diabetesatlas.org* (accessed March 2, 2022).

2. Sherwani, S. I., Khan, H. A., Ekhzaimy, A., Masood, A., & Sakharkar, M. K. (2016). Significance of HbA1c test in diagnosis and prognosis of diabetic patients. Biomarker insights, 11, BMI-S38440.

3. Service, F. J. (2013). Glucose variability. Diabetes, 62(5), 1398-1404.

4. Rodbard, D. (2016). Continuous glucose monitoring: a review of successes, challenges, and opportunities. Diabetes technology & therapeutics, 18(S2), S2-3.

5. Riddle, M. C., Gerstein, H. C., & Cefalu, W. T. (2017). Maturation of CGM and glycemic measurements beyond HbA1c—a turning point in research and clinical decisions. Diabetes Care, 40(12), 1611-1613.

6. Broll, S., Urbanek, J., Buchanan, D., Chun, E., Muschelli, J., Punjabi, N. M., & Gaynanova, I. (2021). Interpreting blood GLUcose data with R package iglu. PloS one, 16(4), e0248560.

7. Rodbard, D. (2009). Interpretation of continuous glucose monitoring data: glycemic variability and quality of glycemic control. Diabetes technology & therapeutics, 11(S1), S-55.

8. Nguyen, M., Han, J., Spanakis, E. K., Kovatchev, B. P., & Klonoff, D. C. (2020). A review of continuous glucose monitoring-based composite metrics for glycemic control. Diabetes technology & therapeutics, 22(8), 613-622.

9. Service, F. J., Molnar, G. D., Rosevear, J. W., Ackerman, E., Gatewood, L. C., & Taylor, W. F. (1970). Mean amplitude of glycemic excursions, a measure of diabetic instability. Diabetes, 19(9), 644-655.

10. Clarke, W., & Kovatchev, B. (2009). Statistical tools to analyze continuous glucose monitor data. Diabetes technology & therapeutics, 11(S1), S-45.

11. Hill, N. R., Hindmarsh, P. C., Stevens, R. J., Stratton, I. M., Levy, J. C., & Matthews, D. R. (2007). A method for assessing quality of control from glucose profiles. Diabetic medicine, 24(7), 753-758.

12. Leelarathna, L., Thabit, H., Wilinska, M. E., Bally, L., Mader, J. K., Pieber, T. R.,… & Hovorka, R. (2020). Evaluating glucose control with a novel composite continuous glucose monitoring index. Journal of diabetes science and technology, 14(2), 277-283.

13. Battelino, T., Danne, T., Bergenstal, R. M., Amiel, S. A., Beck, R., Biester, T., … & Phillip, M. (2019). Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. Diabetes care, 42(8), 1593-1603.

14. Kovatchev, B. P., Otto, E., Cox, D., Gonder-Frederick, L., & Clarke, W. (2006). Evaluation of a new measure of blood glucose variability in diabetes. Diabetes care, 29(11), 2433-2438.

15. Moscardó, V., Giménez, M., Oliver, N., & Hill, N. R. (2020). Updated software for automated assessment of glucose variability and quality of glycemic control in diabetes. Diabetes technology & therapeutics, 22(10), 701-708.

16. Zhang, X. D., Zhang, Z., & Wang, D. (2018). CGManalyzer: an R package for analyzing continuous glucose monitoring studies. Bioinformatics, 34(9), 1609-1611.

17. Vigers, T., Chan, C. L., Snell-Bergeon, J., Bjornstad, P., Zeitler, P. S., Forlenza, G., & Pyle, L. (2019). cgmanalysis: an R package for descriptive analysis of continuous glucose monitor data. PLoS One, 14(10), e0216851.

18. Sechterberger, M. K., Luijf, Y. M., & DeVries, J. H. (2014). Poor agreement of computerized calculators for mean amplitude of glycemic excursions. Diabetes technology & therapeutics, 16(2), 72-75

19. Fernandes, N. J., Nguyen, N., Chun, E., Punjabi, N. M., & Gaynanova, I. (2022). Open-Source Algorithm to Calculate Mean Amplitude of Glycemic Excursions Using Short and Long Moving Averages. Journal of Diabetes Science and Technology, 16(2), 576–577.

20. Piersanti, A., Giurato, F., Burattini, L., Tura, A., & Morettini, M. (2021). Comparison of Software Packages for the Analysis of Continuous Glucose Monitoring Data. In 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA) (pp. 1-6). IEEE.

21. Akturk, H. K., Vigers, T., Forlenza, G., Champakanath, A., & Pyle, L. (2022). Comparison of Cgmanalysis, a Free Open-Source Continuous Glucose Monitoring Data Management and Analysis Software, with Commercially Available CGM Platforms: Data Standardization for Diabetes Technology Research. Diabetes technology & therapeutics, 24(1), 54-60.

22. Dua, D. & Graff, C. (2019). UCI Machine Learning Repository [*http://archive.ics.uci.edu/ml]*. Irvine, CA: University of California, School of Information and Computer Science.

23. Martin, M., Chun, E., Buchanan, D., Bhat, R., Cass, S., Wang, E., Senthil, S. & Gaynanova, I. (2021, April 27). irinagain/Awesome-CGM: List of public CGM datasets (Version v1.1.0). Zenodo. DOI 10.5281/zenodo.3895210

24. Bergenstal, R. M., Beck, R. W., Close, K. L., Grunberger, G., Sacks, D. B., Kowalski, A., ... & Cefalu, W. T. (2018). Glucose management indicator (GMI): a new term for estimating A1C from continuous glucose monitoring. Diabetes care, 41(11), 2275-2280.

25. Fabris, C., Facchinetti, A., Sparacino, G., Zanon, M., Guerra, S., Maran, A., & Cobelli, C. (2014). Glucose variability indices in type 1 diabetes: parsimonious set of indices revealed by sparse principal component analysis. Diabetes technology & therapeutics, 16(10), 644-652.

26. Fabris, C., Facchinetti, A., Fico, G., Sambo, F., Arredondo, M. T., Cobelli, C., & MOSAIC EU Project Consortium. (2016). Parsimonious description of glucose variability in type 2 diabetes by sparse principal component analysis. Journal of diabetes science and technology, 10(1), 119-124. ■

THE IQ CONSORTIUM – CMC AND BIOSTATISTICS STATISTICS LEADERSHIP GROUP

David Christopher (Merck), Erik Talens (Merck), and Phillip Yates (Bristol Myers Squibb)

Introduction

The International Consortium for Innovation and Quality in Pharmaceutical Development (IQ Consortium or IQ, www.iqconsortium.org) is a not-for-profit organization comprised of pharmaceutical and biotechnology companies. The IQ's vision is to be the leading science-based organization for advancing innovative solutions and to enable companies to bring quality medicines to patients. The IQ has close to 40 member companies with over 2,000 annual participants across various activities and celebrated its 10th anniversary in 2020. Member company scientists and engineers can participate on one or more of approximately 100 Working Groups (WG) broadly aligned under four themes: Chemistry, Manufacturing, and Controls (CMC), Life Sciences, Quality, and Statistics. These four areas are subdivided into 10 Leadership Groups (LG), four CMC LGs and four Life Sciences LGs, where both the Quality and Statistics LGs are split between two Forums. Figure 1 illustrates the organizational bridge between the diverse set of WGs (not shown) that operate within each LG and the Board of Directors and labels the LGs by the area of scientific collaboration.

As a science- and technology-based collaborative organization, the IQ has five strategic objectives. First, to collaborate across member companies in a 'N > 1' manner with cross-functional datasets and to form joint scientific positions/conclusions. Second, to advance relationships with other consortia, academic, or government research institutes, e.g., NCATS or PhRMA, to promote scientific excellence and harmonization. The IQ also seeks to proactively engage global regulators either directly, in collaboration with external partners, or via joint meetings on select development topics. Sharing results in the peer-reviewed or trade literature, via presentations or involvement at workshops or symposia, or providing online webinars is the fourth objective. Finally, they seek to ensure the continued value of the IQ through committed leadership, productive collaborations, priority-setting, and talent pool

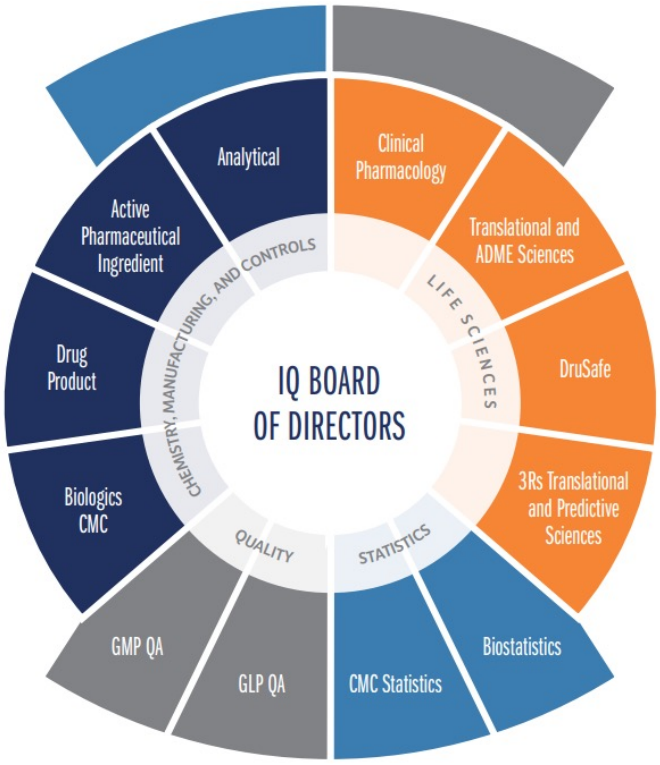


Figure 1: IQ Organization Layout, The IQ has 10 Leadership Groups segmented into four areas (CMC, Life Sciences, Quality, and Statistics) for technical and scientific exchange (graphic courtesy of the IQ Annual Report, iqconsortium.org/about/annual-reports/annual-report-2020). An extensive number of Working Groups, not shown, report to the Leadership Groups.

engagement/leadership succession. The IQ may not be immediately familiar to the at-large biopharmaceutical community; but, a large number of our drug discovery and development subject matter expert colleagues are involved in the IQ.

Two IQ Affiliates, for drug induced liver injury (DILI, *www.iqdili.org*) and microphysiological systems ('organs-on-a-chip', *www.iqmps.org*), also operate in the IQ framework and showcase a tailored scientific pursuit of interest to member companies. The IQ has served as a test bed/launch pad leading to de novo consortia formation, the Allotrope Foundation (*www.allotrope.org*) and the Enabling Technologies Consortium (*www.etconsortium.org*). IQ secretariat,

technical, and legal support is provided by the law firm of Faegre Drinker Biddle & Reath. Secretariat support, in addition to coordinating routine administrative and meeting tasks, provides a valuable infrastructure for effectively sharing and managing both cross-company datasets and communication channels both internal and external to the IQ. Data sharing, industry surveys/benchmarking, and collective positions on scientific topics are just three common examples where the whole exceeds the sum of the parts or the individual contribution of a single member company. In this short introduction to the IQ we highlight how the Statistics Leadership Group (SLG) has contributed to one or more of the IQ's stated objectives.

Statistics Leadership Forums

The majority of the WGs operate in the CMC and Life Sciences areas. The CMC and Biostatistics Forums were created to loosely align activities under these two broad umbrellas. To help provide limited context and reinforce IQ's scientific emphasis here are examples of WGs in each area, respectively: Dissolution (CMC Analytical LG), Nitrosamine (CMC Drug Substance LG), Continuous Manufacturing (CMC Drug Product LG), Subvisible Particles (CMC Biologics CMC LG), Immunogenicity (Life Sciences Clinical Pharmacology LG), Pediatrics PBPK (Life Sciences Translational and ADME Sciences LG), Pre-FIH Tox Attrition (Life Sciences DruSafe LG), and Recovery Animals (Life Sciences 3Rs Translational and Predictive Sciences LG). The Control Strategy Global Harmonization Metrics WG is an example of a Quality LG effort with regulatory dimensions whose title reflects components of the IQ's strategic mission. The CMC and Biostatistics Forums, or SLGs, do not currently have any dedicated WGs but directly support several WGs, a common situation for consulting statisticians where impact is achieved not by being in the limelight but from the trenches. Currently, each SLG has approximately 20-25 statisticians representing approximately 15 member companies. Member company statisticians can have direct involvement with a WG and not actively participate on a SLG. In part, SLG participants are encouraged to act as a liaison between their internal statistician communities and their scientific counterparts. Both technical contributors and managers participate on each of the two Forums.

CMC has traditionally been a focus area for nonclinical statisticians and supported by the CMC Statistics Leadership Group (CMC SLG). Some current CMC SLG members participated on the PhRMA Technical Committee, a precursor to the CMC SLG, more than 2 decades ago. As the value and role of CMC-related statistics evolved the IQ has emerged to offer a part-time 'home' for the CMC statistician community with direct ties to our large and diverse customer base. The Biostatistics Forum or SLG was formed in 2019 to explore and address topics for statisticians supporting domains outside of the CMC space, e.g., drug safety, in vitro/vivo assays. In the paragraphs below we provide a snapshot of some select activities.

CMC – Jyh-Ming Shoung (Janssen), with other member company statisticians, authored an in-depth internal whitepaper on statistical considerations for modeling risk-based predictive accelerated stability studies as part of supporting the Analytical LG. Lori Pfahler, recently retired from Merck, presented at the 2021 Nonclinical Biostatistics Conference on CMC in relation to statistics and data science that was, in part, based on a series of discussions among the CMC SLG. Jose Ramirez, now at Kite Pharma, co-presented a four-part webinar series on computational Bayesian methods for CMC applications that was attended by several hundred viewers across the globe.

Reviewing and offering an opportunity to provide feedback on upcoming regulatory guidances and standards has occurred. Member companies often choose to provide individual feedback via their own internal mechanisms; but, the opportunity to present a collective response has occurred in the IQ framework (e.g., the Dissolution WG's endorsement of the use of apex vessels for USP dissolution evaluations, review of USP <1095> for batch release dose uniformity testing).

The CMC SLG provides a regular and recurring framework for discussing items of mutual interest. For example, it provided an effective forum to consider and debate important issues around limitations of the widespread similarity factor (f2) approach for comparing dissolution profiles. Using IQ resources, the group conducted a member company survey, engaging analytical, regulatory, and statistics functions, to better gauge the impact of its (essentially statistical) limitations. The survey concluded that the f2 approach was a 'nuisance' but not a pain point that warranted engaging resources to address. The IQ framework provided an opportunity for a clearer understanding of both the statistical aspects of this approach as well as practical implications. Stan Altan (Janssen) has been involved in on-going and evolving clinically relevant discussions regarding the 'safe space' concept in relation to bioequivalence. This

developing area relies on a deeper physiology-based pharmacokinetic (PBPK) understanding of the interplay between dissolution and in vivo outcomes. Adopting the 'safe space' concept is an example of a 'patient-centric' implementation of drug product quality. Recent interest in topics such as digital twins or real-world evidence (RWE) have been posed that might otherwise have only been conducted via email or ad hoc at a (virtual) conference. Some have benefitted from CMC SLG feedback on upcoming conference presentations. Unlike the specialization that can occur in the clinical trial space CMC SLG members are often involved across the nonclinical spectrum. Members have recently been involved in the ASA Biopharmaceutical Section's separate Nonclinical Bayesian and p-value workstreams.

Biostatistics – The Biostatistics Forum (Biostatistics SLG) formed with the goal of expanding into areas outside the traditional CMC remit. While expanding to the clinical trial domain was a (currently unrealized) possibility current Biostatistics SLG members work in and discuss matters related to drug safety, aspects of pharmacology, and topics of specific member company interest. For example, several robust discussions resulted from an internal presentation series on anti-drug antibody assay cut point determinations. For statisticians interested in collaborating alongside scientists to improve quantitative rigor and help shape best practices the Life Sciences area is ripe for additional involvement.

Jorge Quiroz and Dave Christopher (both from Merck) engaged with two Translational and ADME Sciences LG (TALG) WGs following their request for SLG involvement. The Bioanalytical Samples QC WG conducted an evaluation to assess agreement for pharmacokinetic data from passed, failed, and retested samples. In a Bioanalysis article involving data from several companies, molecular types, and analytical platforms, the analyses suggested that the bioanalytical methods are very reproducible and that the QC samples improve overall pharmacokinetic assay quality. Separately, a second TALG WG study was conducted to determine the value of duplicate versus singlet-based sampling. In another Bioanalysis article incorporating actual and simulated data from 20 diverse drug candidates across four analytical platforms from eight organizations they proposed that a singlet-based sampling approach is a suitable default for ligand-binding assays whereas a duplicate-based approach is needed where imprecision and/or inaccuracy impedes assay validation. Phillip Yates (BMS) assisted the TALG Induction WG for several years in its evaluation of this preclinical in vitro assay, using shared data generated for this effort by at least seven member companies, as part of the important role it can play in the study of drug-drug interactions. Multiple *Drug Metabolism and Disposition* articles, co-authored by member company statisticians, were published with a view towards influencing regulatory guidelines. The recent FDA guidance on in vitro drug interaction studies, released in early 2020, cites two of these papers as references.

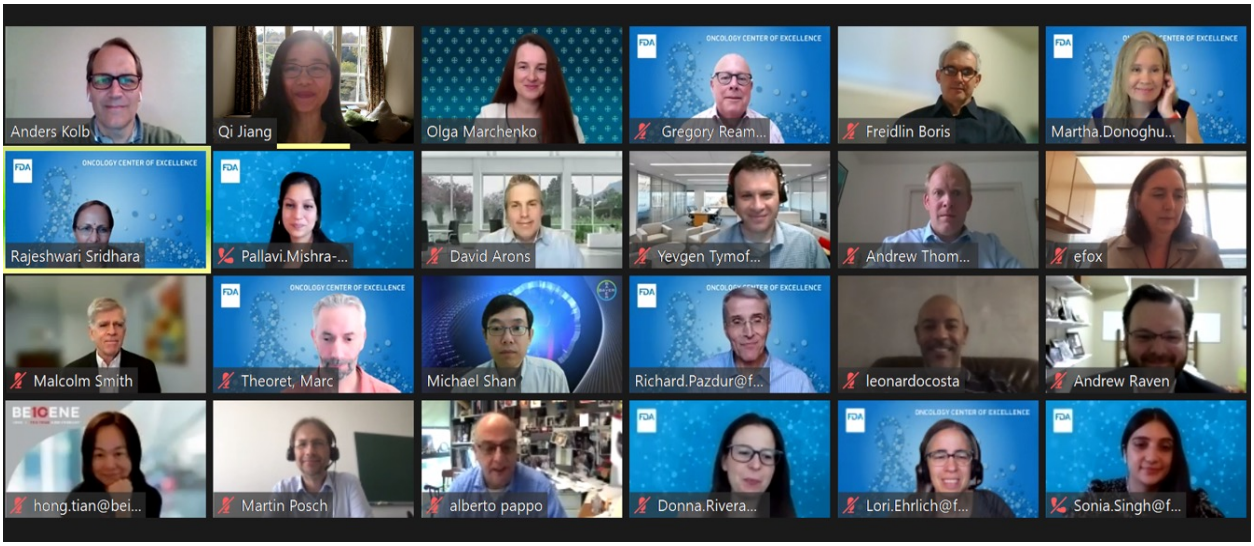
Conclusion

The pre- or nonclinical arena for statisticians is broad and heterogeneous. As highlighted in a recent ASA Biopharmaceutical Report (Fall '21), for US-based colleagues the biennial Nonclinical Biostatistics Conference, Midwest Biopharmaceutical Statistics Workshop, ASA Biopharmaceutical Section nonclinical working group, and the Nonclinical Biostatistics Leaders' Forum are well-known venues for nonclinical statisticians. The IQ, although not statistics-centric per se, is another less-trodden 'home' for nonclinical statisticians with a distinct emphasis on scientific topics not pertaining to clinical trials. A diverse range of scientific and practical topics related to drug discovery and development are encountered and present compelling and challenging problems. The IQ provides, especially for subject matter experts engaged on various WGs, a unique and beneficial platform for cross-company collaboration and another venue for CMC or nonclinical statisticians to contribute and help bring new medicines to patients.

If you would like more information or are employed at a member company and wish to get involved in the IQ please contact Erik Talens (CMC SLG chair, *erik.talens@merck.com*), Phillip Yates (CMC SLG immediate past chair, *phillip.yates@bms.com*), or Dave Christopher (Biostatistics SLG chair, *j.david.christopher@merck.com*). Interested participants are encouraged to refer to a recent Pharmaceutical Technology (*www.e-digitaleditions.com/i/1302985-pharmtech-regulatory-sourcebook-october/59*) article for additional overview or the 2020 annual report (*iqconsortium.org*) to see a select list of publications, comments on regulatory guidances and standards, and assorted presentations. ■

SUMMARY OF ASA BIOP SECTION'S VIRTUAL DISCUSSIONS WITH REGULATORS ON STATISTICAL CONSIDERATIONS IN CLINICAL TRIALS FOR RARE PEDIATRIC CANCERS

Rajeshwari Sridhara (FDA), Olga Marchenko (Bayer), Qi Jiang (Seagen), Elizabeth Barksdale (LUNGevity Foundation), Richard Pazdur (FDA), Gregory Reaman (FDA)



*** Speakers/ Panelists: Dr. Todd Alonzo (COG Group Statistician), Mr. David Arons (National Brain Tumor Society), Dr. Elizabeth Barksdale (LUNGevity Foundation), Dr. Alex Bliu (HC, Canada), Dr. Qiuyi Choo (HAS, Singapore), Dr. Michael Coory (TGA, Australia), Dr. Martha Donoghue (FDA), Lori Ehrlich (FDA), Dr. Leonardo Fabio Costa Filho (ANVISA, Brazil), Dr. Elizabeth Fox (St. Jude), Dr. Boris Freidlin (NCI), Ms. Nancy Goodman (Kids V Cancer), Dr. Doug Hawkins (COG Group Chair), Dr. Qi Jiang, (Seagen), Dr. Pallavi Mishra Kalyani (FDA), Dr. E. Anders (Andy) Kolb (DuPont Hosp. for Children, COG AML Chair), Dr. Helen Mao (HC, Canada), Dr. Olga Marchenko (Bayer), Dr. Eric Ng (Bayer), Dr. Alberto Pappo (Director, Solid Tumor Program at St. Jude), Dr. Richard Pazdur (OCE, FDA), Dr. Dr. Martin Posch (University of Vienna), Mr. Andrew Raven (HC, Canada), Dr. Khadija Rantell (MHRA, UK), Dr. Gregory Reaman (OCE,FDA), Dr. Donna Rivera (OCE, FDA), , Dr. Pourab Roy (FDA), Dr. Sonia Singh (FDA), Dr. Michael Shan (Bayer), Dr. Malcolm Smith (NCI), Dr. Rajeshwari Sridhara (OCE), Dr. Andrew Thompson (EMA), Dr. Hong Tian (BeiGene), Dr. Yevgen Tymofyeyev (J&J), Dr. Jonathon Vallejo (FDA), Jingjing Ye (BeiGene)**

The American Statistical Association (ASA) Biopharmaceutical Section (BIOP) and LUNGevity Foundation hosted open forum virtual discussions on June 24, 2021 with participation from biostatisticians, clinicians, and regulators in a series conducted under the guidance of the U.S. FDA Oncology Center of Excellence's Project SignifiCanT (Statistics in Cancer Trials). The goal of Project SignifiCanT is to advance cancer drug development through collaboration and engagement among stakeholders in the design and analysis of cancer clinical trials. These discussions were organized jointly by the ASA BIOP Statistical Methods in Oncology Scientific Working Group, the FDA Oncology Center of Excellence (OCE), and LUNGevity Foundation.

Randomized clinical trials remain the best method to assess the benefit/risk of investigational treatments. However, all pediatric cancers are rare and molecular characterization of specific cancers has resulted in

even smaller population subsets. In many cases limited sample sizes often make it infeasible to conduct large randomized clinical trials in a timely manner with adequate power and control of type I error rate at a two-sided 0.05 level. To accelerate pediatric cancer drug development, there is a need to identify potential innovative clinical trial design options to evaluate treatments for pediatric cancers with small sample sizes in which randomized trials are not possible. This virtual open forum discussion focused on clinical trial design considerations for evaluating new treatments for pediatric cancers where the standard or traditional randomized trials are deemed infeasible.

The speakers/panelists* for the discussion included members of the BIOP Statistical Methods in Oncology Scientific Working Group representing pharmaceutical companies, representatives from International Regulatory Agencies (FDA, EMA, HC, TGA, ANVISA, MHRA, HSA), clinical investigators, academicians, patient advocacy groups and expert statisticians in industry. In addition, over 100 participants attended the virtual meeting, including representatives from other International Regulatory Agencies (SMC, PMDA, Israel). The discussions were moderated by the BIOP Statistical Methods in Oncology Scientific Working Group co-chairs, Dr. Qi Jiang from Seagen and Dr. Olga Marchenko from Bayer; Dr. Elizabeth Barksdale from LUNGevity Foundation; and Dr. Rajeshwari Sridhara, contractor from OCE, FDA.

This open forum was divided into three sections, each featuring a short presentation followed by a panel discussion, on the following topics: (1) Use of external controls, (2) Frequentist approach, and (3) Bayesian approach. After the introductory remarks by the OCE leadership who highlighted the need for a multi-disciplinary approach to facilitate pediatric drug development, an FDA representative presented regulatory and statistical considerations for externally controlled clinical trials. The presentation acknowledged that while randomized controlled trials are the preferred approach for generating evidence of drug safety and efficacy, in certain cases where randomization is not feasible, an external control arm may be an option for estimating comparative treatment effect, when the clinical course of the disease is well understood. Furthermore, to consider external control as a comparator, the data must be fit for purpose with appropriate quality and completeness, including clear definitions of disease criteria, baseline demographics and disease characteristics, key clinical covariates, endpoints, standardized measurements, timing and follow-up schedule of assessments, temporality of data, and access to patient-level data are necessary. The panel members pointed out that within currently identified molecular subgroups, external control data may be less than optimal, jeopardizing their use. It was recognized, however, that the external and relevant data can provide a benchmark and be used to generate hypotheses for future clinical trials. In some cases, there may no longer be equipoise when the adult data are already available, and drug is shown to be effective in adults for a disease and/or biomarker defined disease that also occurs in children, or when there are data from off label use in children. In general, patients appreciate that using external data can attenuate the chance of being randomized to control. Additionally, use of external control data may reduce the time and cost associated with conducting trials. Minimizing bias is the key issue in considering external controls. Panel members from multiple regulatory agencies stressed the importance of data quality, evaluation in the presence of missing data, overall comparability, and prospective planning in consultation with regulatory bodies.

The second section of the discussion focused on frequentist randomized controlled clinical trials with innovative designs. Presenting for this section, a National Cancer Institute (NCI) statistician proposed a strategy of relaxing the evidentiary threshold (Type I error rate) in randomized controlled clinical trials for pediatric cancers, highlighting the importance of randomization even in small populations with limited numbers of patients to participate in clinical trials. The proposal was to consider feasibility-driven evidentiary standards, based on the time required to enroll patients (e.g., 4-5 years may be appropriate for some newly diagnosed patient populations). Other innovative designs, such as use of master protocols with common control arms, can also be efficient ways of conducting the clinical trials. During the discussion, panel members agreed that with the Type I error controlled at one-sided 0.025 level, clinical trials- in particular those for rare pediatric cancers- could take 8-10 years or longer to complete, by which time the drug being investigated may no longer

be reflective of clinical practice. When there is equipoise, small randomized clinical trials are feasible by increasing Type I error, or if the drugs are very effective. A systematic decision approach may be needed, considering the risks of erroneous decisions, limited sample size, and the potential harm of approving a drug which may have limited or no efficacy. It was also noted that the meaningful benefit and risk assessments are different in adults and pediatric patients. Regulators pointed out that there should be flexibility in regulatory considerations depending on the clinical relevance, currently available treatment options, and the magnitude of treatment effect. Some panelists opined that low-powered randomized controlled trials are preferred, compared to uncontrolled single arm trials, with or without use of external control comparator data.

The third section of the discussion focused on use of Bayesian approaches. An industry statistician presented statistical considerations when using a Bayesian design and analysis, particularly understanding the uncertainty of extending or borrowing adult data in evaluating a treatment effect in pediatric patients. Different Bayesian methods such as use of power prior, commensurate prior, and meta-analytic-predictive approach are available. Key considerations include what amount of information can be borrowed from historical/adult/external data and how the operating characteristics (e.g., Type I error) can be controlled. Some of the challenges include concerns regarding simulated Type I error and degree of historical data included in the estimation of treatment effect, recognizing that the amount borrowed is a random variable. In the panel discussion it was acknowledged that efficient trials to test and prioritize drugs are needed, particularly with the development of molecularly targeted therapies. Bayesian designs are attractive for this purpose. However, there are uncertainties in applicability of adult data to children in some indications, including host characteristics and the disease characteristics, preferred outcomes, and feasibility of measurement. As an example, a Bayesian design (AGILE, *https://clinicaltrials.gov/ct2/show/NCT03970447*) with three arms is currently being used in adults with glioblastoma multiforme. Acceptability of Bayesian designs that borrow adult/historical data depends on the availability and quality of data, keeping in mind that children are not young adults with respect to metabolism, risk of toxicity, degree of tolerability, and other characteristics. Most often there is not sufficient prior data to initiate a Bayesian design. Whether Bayesian design is more appropriate than frequentist approach needs to be considered on a case-by-case basis. By borrowing data Type I error will increase and there is potential bias in selecting prior data. Timely discussions with regulators are also critical during the planning stage.

This forum provided an opportunity to have open scientific discussions among a diverse interdisciplinary stakeholder group – clinicians, statisticians, patient advocates, international regulators, and representatives from pharmaceutical companies focusing on emerging statistical issues in cancer drug development. We plan to continue with similar multidisciplinary open forum discussions in the future on a variety of important topics that include statistical aspects in cancer drug development with various stakeholders participation.

Acknowledgement: Authors thank Joan Todd (FDA) and Rick Peterson (ASA) for supporting the forum, and Dr. Jun Yin (Mayo Clinic) for taking the meeting minutes. ■

SPECIAL ISSUES IN THE *JOURNAL OF BIOPHARMACEUTICAL STATISTICS*

Margaret Gamalo (Pfizer; Editor-in-Chief JBS)

Recently, the *Journal of Biopharmaceutical Statistics (JBS)* launched several special issues to create coherent and thematic set of statistical research that can serve as desktop reference for statisticians in the biopharmaceutical industry. The themes were carefully selected to reflect topics that are relevant to the emerging drug development landscape. The special issue may also come from recently concluded conference or symposium discussing thematic statistical research, emerging insights and opinions, and technological trends in the biopharmaceutical industry. Because they have a separate micro-editorial system, special issues leverage networks of key opinion leaders in the field for contributed manuscripts. They also are more expedient because the manuscripts do not go through the usual review process within the JBS system. Generally, the submitting authors in a special issue create an ecosystem of reciprocal peer review. The current list of special issues is described below and include ways to get connected with the respective Guest Editors and whether they are still accepting manuscripts.

The first special issue launched is on **Real World Evidence.** This issue will feature innovative statistical methodologies and case studies that incorporate real-world data (RWD) or RWE into medical product development and regulatory decision-making, especially in areas of unmet medical need. For instance, in clinical trials for treatment of rare diseases or rare, genetically targeted subsets of common diseases, randomization may not be feasible due to timeline, budget, and ethical concerns. Using external controls by borrowing information from RWD sources or historical trials can provide relevant evidence (i.e., RWE) to examine the treatment effect from single-arm trials. Further, innovative adaptive designs enable trialists to utilize historical information from adults or children of other age groups in their planning for pediatric trials, provided that these reference data gives the basis and prior knowledge for extrapolation and inferences of the treatment effect. With the advances of technology and increasingly available RWD/RWE, there is more room for improvement in clinical trial designs and data analysis strategies. The guest editors in this issue are **Junjing Lin (Takeda)** and **Helen Li (BMS)** and came up with an issue that presents new ideas in this field, e.g., methodological papers such as entropy balancing (Yu et al. 2022), small-sized trial (Jiao et al. 2022), resampling approaches for estimating similarity measure (Li et al. 2022), sampling importance resampling method for borrowing (Sachdeva et al. 2022), propensity score based prior construction (Baron et al. 2022, Lu et al. 2021), deep learning (Zhan et al. 2021), combining survival data reconstruction technique with balancing weights (Getz et al. 2021), and intermediate outcome assisted borrowing (Liu et al. 2022). Additionally, a comprehensive review of propensity score methods under Bayesian framework was provided by Lin and Lin (2021), and extensive simulation studies on different propensity score adjustment methods with various priors were conducted by Wang et al (2022); a case example of propensity score matching was illustrated by Yin et al. (2022). The issue should be published very soon.

Another special issue that will be published subsequently is on the conference for the **2021 Duke Industry Statistics Symposium (DISS)** around the theme of "*Emerging Clinical Initiatives in Pharmaceutical Development*". The DISS is organized by the Department of Biostatistics and Bioinformatics, Duke University School of Medicine. It was established 7 years ago to discuss challenging issues and recent advances related to the clinical development of drugs and devices and to promote research and collaboration among statisticians from industry, academia, and regulatory agencies. The Annual meeting happened on April 21-23, 2021. The editors for this special issue are **Herbert Pang (Genentech)** and **Nelson Lu (FDA)** Some of the innovative research and clinical trial strategies that will be coming out of this issue are group sequential design for randomized clinical trials with adaptive information borrowing, propensity score-integrated approach to

survival analysis, patient-centric clinical development, randomization tests with multiple imputation for handling missing data, adaptive external evidence incorporation in sequentially monitored clinical trials, and data management of future digital clinical trials.

The **Estimand** issue features manuscripts relevant to the discussion of estimands and associated sensitivity analysis in a broad context, i.e., not just for efficacy but also for safety and benefit-risk evaluation, in pre-marketing and post-marketing settings. ICH E9 (R1) states that "An estimand is a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective". In drug development, the estimand of interest might be different for patients, regulators and payors. Carefully understanding the estimand of interest, the intercurrent events and the associated sensitivity analysis is therefore important. Due to this complexity, the statistical community needs to take a lead here in raising awareness and understanding of the impact of using different estimands on the interpretation and decision-making processes in drug development. In light of this, the *Journal of Biopharmaceutical Statistics* decided to dedicate a special issue to this topic even if there has been so much attention on estimands in that past 3 years. This issue is edited by **Bill Wang (Merck)** and **Yodit Seifu (BMS)** and it will contain topics that also highlight other avenues where estimands have not been common, e.g., assessing adverse events in clinical trials, estimands in benefit-risk determinations, estimands in vaccine trials, and in associated causal estimands for hybrid trials. The special issue is slated to be published toward the end of this year.

Another issue that will be published toward the end of this year is on the conference proceeding for the 2021 International Chinese Statistical Association (ICSA) Applied Statistics Symposium which was held on September 12-15, 2021. The theme of this conference is Leading with Statistics and Innovation and had a primary audience among statisticians working in academia, government, and industry. The guest editors are **Guoqing Diao (GWU), Judy Wang (GWU), Qing Pan (GWU), Lu Mao (Wisconsin)** and **Junjing Lin (Takeda).** This issue will showcase research on robust estimates of regional treatment effects with ordinal responses in multiregional RCTs, use of surrogate information to improve confirmatory platform trial with sample size re-estimation, indirect comparison with real-world data for the survival endpoint under non-proportional hazards, Bayesian design for platform trials with multiple endpoints, multi-stage dose expansion cohort design with Bayesian stopping rule.

There are three special issues that are still accepting contributed manuscripts. The first one is on **Statistics in Pediatric Drug Development.** This issue discusses theory, applications, and practical considerations when using innovative statistical methodologies or clinical trial designs to support new therapies in pediatrics. It will include, but is not limited to, the following topics

- extrapolation from adults to pediatrics, older age group to younger age group, etc.
- practical statistical considerations in the analysis and reporting of results of pediatric clinical trials,
- regulatory and statistical issues encountered in the design and analysis of pediatric clinical trials,
- statistical methodologies in demonstrating dose-exposure-response similarity, and
- statistical methods for the assessment of both limited short- and long-term safety data
- real world databases in pediatrics and their use in regulatory decision making (RWE)
- Innovative design and analysis methods in pediatric rare diseases
- pediatric epidemiology considerations in evaluating drug efficacy, safety, and effectiveness

The guest editors are **Jingjing Ye (BeiGene), Rima Izem (Novartis),** and **YJ Choi (Genentech)**, and **Margaret Gamalo (Pfizer).** There has been a lot of interest and numerous inquiries to this issue - a testament to the fact that the pediatric drug development space needs innovative and efficient statistical methodologies to close the lag time between labelling in adults and labelling in children. Contributors are encouraged to reach out to any of the guest editors for details on the submission.

The next special issue is on **Advances in statistical methods for the assessment and validation of clinician, patient-reported, digital health, and photography outcomes**. JBS recognizes that the industry is undergoing a shift toward implementing a paradigm that accentuates the influence of key stakeholders like

patients, regulators, and payers. For example, clinical outcomes with patient-oriented clinical relevancy and impact on quality of life describing added dimensions of therapeutic value are taken into consideration when designing pivotal studies and evaluating the added benefit of medicines. Increased emphasis has been given to the validity, reliability, responsiveness, and interpretation of patient-centric outcomes. As such, psychometric methodology is continually evolving with a directive for developing and refining more effective and meaningful outcomes.

In addition, the recent explosion of data collected from digital health technologies (DHTs), including wearable sensors and connected devices, have captured human behaviors such as physical activity, physiology, the environment and the broader functional status of the patients – which are advancing patient outcomes even further. DHTs are enabling the adoption of decentralized clinical trials and engaging patients and collecting data in their communities and homes using remote technology more frequently and unobtrusively. This growing deployment of DHTs and other digital tools such as photography in pharmaceutical research, and the shift to more community-based trials, present new challenges to the derivation, statistical analysis and interpretation of clinical outcomes and digital health measures.

For patient-centric and DHS data to support conclusions on efficacy and safety profile of a therapeutic product, regulatory and reimbursement agencies require that these are well-defined, accurate, sound, and clinically meaningful to patients in assessing their symptoms and how they feel and function.

For this reason, the *Journal of Biopharmaceutical Statistics* invites authors to submit papers for a special issue on Advances in Statistical Methods for the Assessment and Validation of Clinician, Patient-Reported, Digital Health, and Photography Outcomes. This special issue will feature innovative statistical methodologies, novel study designs, instructive case studies, and practical conceptual considerations in the design, validation, analysis, interpretation, and reporting of clinical outcome, digital health, and photographic assessments. The guest editors are **Joseph Cappelleri (Pfizer) Charmaine Demanuele (Pfizer), Jessica Roydhouse (U Tasmania),** and **Margaret Gamalo (Pfizer).**

The last in the current list of special issues is on Insights into drug development in the Emerging Markets: Statistical considerations and opportunities. "Emerging markets" is a frontier representing an exceptional opportunity for the biopharmaceutical industry as governments in these regions are looking to reform public healthcare and grant easier access to innovative medicine. Gaining insights into the current and anticipated future market trends is key to navigating any emerging market and proper integration of their needs into the global drug development. This includes gathering insights on local healthcare access, care paths, reimbursement, and funding (including for both original and generic products), health policies, market demands, the competitive landscape, and the data requirements for market access. There are many development issues including efficacy and dosing reflecting variations in a drug's efficacy and toxicity among different geographical or ethnic populations; understanding physicians' and patients' preferences and responding to them effectively is important to success in emerging markets; fixed-dose combination medicines—drug therapies with two or more active pharmaceutical ingredients combined in one tablet—are preferred in many markets, but their popularity varies greatly from one to the other; prerequisite local presence differs in China, India, South Korea, and Taiwan for local product registrations.

Due to the prominence of this topic, the special issue will feature statistical methodologies on topics that are relevant to the emerging markets, case studies, practical statistical considerations in designing trials that include emerging markets, issues encountered in the analysis of data supporting efficacy and safety in the emerging markets, statistical insights into post marketing activities and, in general, opportunities for further research and integration. For further questions regarding submission, please contact one of the Guest Editors: **Yuki Ando (PMDA), Joseph Beyene (McMaster) Jie Chen (Overland), Margaret Gamalo (Pfizer), Mahesh Iyer (Parexel), Shoichi Ohwada (Daiichi-Sankyo), Hsiao-Hui Tsou (National Health Research Institutes), Jun Wang (NMPA),** and **Ying Wu (Southern Medical U)**.

There will be other special issues that will be announced in the near future. For those who are interested in becoming a guest editor, please send me an email at *margaret.gamalo@pfizer.com* and we can discuss potential topics and its feasibility. There are

topics that does not even have to reflect current statistical developments but something that will benefit statisticians within the biopharmaceutical industry. The journal is also keen on partnerships with conferences and workshops so we can push innovative research and insights in the most expedient manner. JBS also encourages uncommon foresight reflecting vision of where the industry is heading or where we as statisticians want it to be.

References

Yu, G., Bian, Y., Gamalo-Siebers, M. (2022) Power Priors with Entropy Balancing Weights in Data Augmentation of Partially Controlled Randomized Trials. *Journal of Biopharmaceutical Statistics,* DOI: *10.1080/10543406.2021.2021226.* Epub ahead of print. PMID: 35072583.

Jiao, F., Chen, Y.F., Min, M., Jimenez, S. (2022) Challenges and potential strategies utilizing external data for efficacy evaluation in small-sized clinical trials, *Journal of Biopharmaceutical Statistics,* DOI: 10.1080/10543406.2021.2011906 . Epub ahead of print. PMID: 34986063.

Li, H., Li, Q., Tiwari, R. (2022) Borrowing External Data to Establish A Hybrid Control Arm in Randomized Clinical Trials. *Journal of Biopharmaceutical Statistics,* DOI: *10.1080/10543406.2021.2021227.* Epub ahead of print. PMID: 35067183.

Sachdeva, A., Guha, S, Tiwari, R. (2022) A Novel Approach to Augment Single-arm Clinical Studies with Real-World Data. *J Biopharm Stat.* DOI: 10.1080/10543406.2021.2011902. Epub ahead of print. PMID: 34958629.

Baron, E., Chen, M.H., Zhu, J., Tang, R. (2022) Bayesian Divide-and-Conquer Propensity Score Based Approaches for Leveraging Real World Data in Single Arm Clinical Trials. *J Biopharm Stat.* DOI: 10.1080/10543406.2021.2011904. Epub ahead of print. PMID: 35094643.

Zhan, T., Zhou, Y., Geng, Z., Gu, Y., Kang, J., Wang, L., Huang, X., Slate, E.H. (2022) Deep historical borrowing framework to prospectively and simultaneously synthesize control information in confirmatory clinical trials with multiple endpoints. *J Biopharm Stat.* 2021 Oct 10:1-17. DOI: 10.1080/10543406.2021.1975128.

Liu, Y., Lu, B., Foster, R., Zhang, Y., Zhong, Z.J., Chen, M.H., Sun, P. (2022) Matching design for augmenting the control arm of a randomized controlled trial using real-world data. *J Biopharm Stat.* DOI: 10.1080/10543406.2021.2011900. Epub ahead of print. PMID: 35098880.

Lu, N., Wang, C., Chen, W.C., Li, H., Song, C, Tiwari, R., Xu, Y., Yue, L.Q. (2022) Propensity score-integrated power prior approach for augmenting the control arm of a randomized controlled trial by incorporating multiple external data sources. *J Biopharm Stat.* 2021 Nov 10:1-12.DOI: 10.1080/10543406.2021.1998098.

Getz, K., Mamtani, R., Hubbard, R.A. (2022) Integrating real world data and clinical trial results using survival data reconstruction and marginal moment-balancing weights. *J Biopharm Stat.* 2021 Nov 10:1-13. DOI: 10.1080/10543406.2021.1998097.

Lin, J., Lin, J. (2022) Incorporating propensity scores for evidence synthesis under bayesian framework: review and recommendations for clinical studies. *J Biopharm Stat.* 2021 May 16:1-22. DOI: 10.1080/10543406.2021.1882481.

Wang, X., Li, X., Jemielita, T., Suttner, L. (2022) Propensity score-integrated Bayesian prior approaches for augmented control designs: A simulation study. *J Biopharm Stat*. DOI: 10.1080/10543406.2021.2011743. Epub ahead of print. PMID: 34939894.

Yin, X., Davi, R., Mishra-Kalyani, P., Stuart, E.A., Stewart, M., Sridhara, R., Dawson, J. (2022) Exploring the Potential of External Control Arms created from Patient Level Data: A case study in non-small cell lung cancer. *J Biopharm Stat*. DOI: 10.1080/10543406.2021.2011901. Epub ahead of print. PMID: 34986069. ■

UPCOMING CONFERENCES

2022 ASA Biopharmaceutical Section Regulartory-Industry Statistics Workshop

The ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop is sponsored by the ASA Biopharmaceutical Section in cooperation with the FDA Statistical Association. The conference lasts three days (September 20, 2022 – September 22, 2022), with invited sessions co-chaired by statisticians from industry, academia, and the FDA and short courses on related topics offered on the first day of the workshop. To find out more visit: *https://ww2.amstat.org/meetings/biop/2022/*
Conference Early Registration Opens: June 15, 2022 and Early Conference Registration Closes: August 17, 2022

2022 WNAR

The WNAR 2022 will be held virtually from June 10-15, 2022. The meeting brings together researchers and practitioners from academia, industry and government, connected through a common interest in Biometry. Short Courses will be held from June 10-11, 2022 and invited Oral Sessions, Contributed Oral Sessions, Student Paper Sessions will be held from June 13-15, 2022. Conference registration is free for all student members of IBS. To register visit here: *https://wnarofibs.wildapricot.org/WNAR2022*

2022 Symposium on Data Science & Statistics (SDSS)

The ASA's fifth annual SDSS will be held in Pittsburgh, PA from June 7–10, 2022. SDSS provides a unique opportunity for data scientists, computer scientists, and statisticians to come together and exchange ideas. To register visit here: *https://ww2.amstat.org/meetings/sdss/2022/registration.cfm*. You may register online or using the PDF form. For the latter, the deadline is May 31, 2022.

JSM 2022

Joint Statistical Meetings (JSM) is the largest gathering of statisticians and data scientists held in North America. This year's meeting will be held between August 6, 2022 and August 11, 2022. It is also one of the broadest, with topics ranging from statistical applications to methodology and theory to the expanding boundaries of statistics, such as analytics and data science. JSM also offers a unique opportunity for statisticians in academia, industry, and government to exchange ideas and explore opportunities for collaboration. Early registration opens May 2, 2022 and closes on May 31, 2022. Regular registration is from June 1, 2022 to June 30, 2022. Late registration starts on July 1, 2022. To register visit here: *https://ww2.amstat.org/meetings/jsm/2022/registration.cfm* ■