Kammerman
CONSULTING

# Summary:
# 2019 DIA/FDA Biostatistics Industry and Regulator Forum

This document summarizes key messages from the DIA/FDA Biostatistics Industry and Regulator Forum. The forum was held 8 April 2019 through 10 April 2019 at the Bethesda North Marriott Hotel and Conference Center in North Bethesda, Maryland, USA.

The document is organized in 3 sections:

1. **Links to forum information** contains URLs that will take you to the official forum website and agenda.
2. **Key messages**. These are our high-level takeaways. If you want more information, the hyperlinks will take you to detailed notes.
3. **Dr. Kammerman's notes** contains the notes she captured from the sessions that she attended.

## 1. LINKS TO FORUM INFORMATION:

- Website for 2019 DIA/FDA Biostatistics Industry and Regulator Forum
- Forum overview, highlights and agenda (PDF)
- Online agenda - detailed

## 2. KEY MESSAGES

- Dr. Janet Woodcock introduced the proposed reorganization of the Office of New Drugs (OND)/CDER/FDA, which is designed to meet current and future scientific needs. New review processes will be issue-based. Dr. Woodcock described implications for the statistical community. Her keynote address is summarized here.
- The session on complex innovative trial designs updated the audience on FDA's pilot meeting program. Notably, the public can submit questions online and FDA will post responses. Presenters identified key operating characteristics that need to be targeted by simulations, described the concepts of a simulation plan and simulation report, and steps needed to reduce the introduction of bias into the conduct of a complex innovative trial. Notably, FDA plans to develop guidelines for simulation software.
- How to communicate with non-statisticians -- especially clinicians, and how to educate them was a key theme of the Bayesian methods session. I liked Dr. Wathan's use of color codes in a graphic that compared operating characteristics obtained from frequentist and Bayesian approaches.
- Real world data and big data were themes throughout the meeting and across many sessions, three of which focused entirely on RWD and RWE.
    - One session discussed the role of pragmatic trials in the regulatory setting and considerations in NDA and BLA submissions, especially for rare diseases and oncology. FDA has a website dedicated to RWD/RWE. The Q&A was extremely lively, with commentators saying the FDA needs to catch-up and stay current with machine-learning and artificial intelligence. The potential for RWD studies that mimic RCTs and a path to how they could eventually be accepted by regulators was also discussed.

  - The senior leaders town hall discussed the role of big data and the need for statisticians to develop new skills, potentially by partnering with and learning from, for example, computer scientists and machine learners. Pharmaceutical companies are transforming into data organizations and statisticians will need to become more comfortable with research directions, rather than focusing on research questions, in order to uncover insights.
  - A session on causality, artificial intelligence and big data stressed the need for a transparent representation of the mechanisms underlying the phenomenon under investigation. Dr. Bareinboim led an informative discussion on graphical views of causal inference with an example that used observational data. He recommended a recently published book that reviews causal inference over the last 40 years. The role of targeted learning (vs. machine learning) in generating estimates from real world data was another topic.
- The session on patient-focused drug development mentioned FDA's PRO guidance will be re-written. I welcomed the comment that biostatisticians and psychometricians are not the same profession, nor do they have the same skills.
- Speakers in the complex generic drugs session discussed statistical challenges in the development of generic drugs, especially for complex generics. The speakers defined a complex generic, which was introduced in GDUFA II, and addressed clinical endpoints, non-inferiority, and when PK studies are needed. Complex generics offer statistical opportunities to develop innovative bioequivalence methods. Of note, the FDA has many product-specific guidance documents for the development of generics (https://www.accessdata.fda.gov/scripts/cder/psg/index.cfm).
- A theme of FDA's openness to statistical innovation wove its way through many sessions. The master protocol session emphasized this theme and pointed to an NEJM article by Drs. Woodcock and LaVange on master protocols and the pilot program for complex innovative trial designs. Dr. Berry described an approach to estimating efficacy between treatment arms when the standard of care has changed over time. The history of all enrolled patients and the overlap of multiple arms can be used to derive these estimates.
- The session on cancer patient journeys and estimands contained many intriguing topics, starting with how to handle intercurrent events that are unique to oncology studies. Because many study endpoints are not what matter most to patients, speakers emphasized the need to communicate 'patient-friendly' results, in addition to traditional endpoints such as PFS, ORR, and OS. One speaker gave examples of how to handle intercurrent events and their associated estimands, and mentioned an oncology estimands working group with members from 19 companies. I was fascinated by the CAR-T in lymphoma case study, in which some subjects withdrew from the study before the start of their infusion. Reasons included deterioration from baseline, improvement from baseline and no longer required CAR-T, or CAR-T manufacturing failed. The EMA and FDA approaches differed on the scientific question of interest, with EMA focusing on patients enrolled and FDA focusing on those with disease at time of infusion. The panelists included Dr. Temple and a discussion of when in development to start defining estimands, if estimands are appearing in protocols and how to communicate results to patients.
- In the rare disease session, speakers emphasized that registries are typically useful for designing studies and not usually a source of historical control data. Natural history data are needed to design SMART studies. The need to reduce variability within small studies was stressed with an example of taking multiple blood pressure readings instead of a single reading. Networking sessions with HCPs and families are an avenue for understanding the practicalities of a disease and conducting a study, and identifying biomarkers and other outcomes that will lead to higher quality studies.

## 3. DR. KAMMERMAN'S NOTES

Links to Notes for Keynote Address and Sessions:

Keynote Address: Modernizing the New Drugs Regulatory Program: Rationale and Progress

Session 1: Promoting Complex Innovative Trial Designs: An Update and Discussion on the Pilot Meeting Program

Session 2: Challenges and Opportunities for Bayesian Methods in Confirmatory Setting

Session 3: Real World Data and its Fit for Regulatory Purpose

Session 4: Advancing Patient Focused Drug Development

Session 5: Innovative Methods for Assessment of Complex Generic Products

Session 6: Master Protocols: Design Considerations and Operational Examples (Case-Studies)

Session 7: Cancer Patient Journeys and Estimands

Session 8: Senior Leaders Town Hall: Bridging Statistical Science with Data Science in Drug Development

Session 9: Unhealthy Safety Assessment: Moving Towards Better Characterization of Patient Harms

Session 10: Opportunities for Statistical Leadership and Innovations in Rare Disease Therapies

Session 11: Causality, Artificial Intelligence and Big Data

### KEYNOTE ADDRESS: MODERNIZING THE NEW DRUGS REGULATORY PROGRAM: RATIONALE AND PROGRESS

JANET WOODCOCK, MD, DIRECTOR, CENTER FOR DRUG EVALUATION AND RESEARCH, FDA

The Office of New Drugs (OND) is being re-organized to meet current scientific needs. The current state of regulatory science faced by OND requires modernization of the new drug program to stay current. Dr. Woodcock described the history of the 1962 requirements for efficacy and the evolution since then of study designs, interventions, understanding of pharmacology, patient centricity, digital data and the quantity of information now available.

Proposed review processes and reorganization:

- New approach to BLD and NDA reviews
  - Multi-disciplinary team assessment, rather than individual reviews
  - Reviews will be issue-based
    - Will develop list of issues identified during IND submissions and reviews
    - Will allow review process to hit ground running when NDA/BLA submitted
  - Assessments

        - Clinical data scientists to be hired – will do exploratory analyses currently done by medical staff
        - Will develop standardized queries for exploratory analyses – lump/split for safety is a challenge today
        - Statisticians – will continue to evaluate statistical aspects of NDA/BLA
    - Review process
        - Long, repetitive review documents with cut/paste of industry submissions to be phased out and replaced with issue-based reviews
        - Issues will be identified at the beginning of the submission. Will allow more time for mid-cycle reviews, PMRs, finalizing actions, etc.
        - How issues will be resolved
            - Identify who will work on issues.
            - The resolution of the issues will be documented
            - Will provide greater clarity.

    - Assessment (i.e., review) document
        - Contents:
            - Benefit/risk
            - Short reg history, drug development history
            - Issues identified and their resolution
        - Will increase communication transparency. Survey showed outside world cited difficulties finding content in documents.
        - Will support knowledge management – put issues and resolutions in the center
        - Advisory committees will appreciate having the issues laid out – clinical, stats, pharmtox, etc.
- Implementation of new review process
    - Staged implementation
        - Start with a few submissions
        - Training
            - Working together as a team
            - Team-based work products
            - How to involve supervisors without writing a re-review.
            - Get involved in issues
            - How to lead a team
            - How to supervise matrix employees – new for CDER
        - Anticipate new process to be fully in place by end of year
    - Expectations
        - Higher quality
        - More uniform work product
        - Fewer last minute changes
        - Put issues upfront and involve management from the start to solve issues
        - Get work going right away
        - Will give time to negotiate REMS, PMRs, PMCs. Give better product. Finish in timely manner

<u>Proposed reorganization:</u>

- Will be disease-based
    - Rationale

        - Rise of patient advocacy – want to know someone is there for them and their disease
        - OND divisions were arranged around workload
        - Develop flatter organization to accommodate new INDs, etc.
    - Some of the new divisions:
        - Neuroscience - rise in psychiatric treatments. A drought for 40 years. A new increase in conditions: (e.g., postpartum depression), neurodegenerative conditions. Most have no treatments. Pain, addiction medicine
        - Inflammation/immunology: rheumatic, GI, pulmonary, transplants, . Treatments with immuno-suppressants
        - Infectious disease
        - Internal med – endocrine (e.g., diabetes)
        - Compounding
        - Ophthalmology
        - OTC
        - Human reproduction – urology, IVF, peds, maternal and child health, orphan genetic diseases
        - Oncology – hematology will be removed.

    - Better branding to outside world.
    - Stats, clin pharm will realign to serve these disease areas

- Project management will be centralized
    - Want consistent processes.
    - Will help within OND and industry.
    - Will be able to place automation for standard work processes.
- A new policy group – will handle increasing numbers of guidance documents
- Safety
    - Shorter- and longer-term reform for managing post marketing safety issues in a standardized process
    - Multi-disciplinary teams for safety this year. Developing a long-term plan

Here is a summary of Dr. Woodcock's key points and rationale supporting the OND reorg proposal.

- Evolution since the 1962 Kefauver Harris Amendment or "Drug Efficacy Amendment" to the Food, Drug and Cosmetic Act.
    - Study designs (e.g., non-inferiority, adaptive designs)
    - Biology and disease understanding
    - Interventions
        - Anti-sense nucleosides
        - New platforms
        - Gene therapy
        - Cell-based
        - Others we can't imagine
    - Digital revolution
        - Paper to electronic
        - New data sources
        - Genomic data – genetic sequencing

- New understanding of pharmacology
- Drug costs are a challenge
  - Greater evaluation of value of interventions
  - Many patients can't access meds because of costs
  - Society focuses on value and cost
- Expect/encourage patient advocacy
  - What was unusual about AIDS is now common
  - Patient-focused drug development
    - PROs
    - Understanding what disease means to patient
    - Understanding what intervention means to patient
- Increased demand of transparency in regulatory processes.
  - FDA now shares publicly what wasn't shared previously (e.g., safety).
  - Trend to continue
- Industry renewed effort on innovation.
- Increased volume of enterprise
  - Types of companies: small companies, start-ups, international
  - More info coming from industry to FDA. Lots more data submitted for review

*Dr. Woodcock's conclusion: regulatory environment needs to evolve to keep up.*

<u>What the reorganization means for the statistical community:</u>

- Improvements in sharing across teams. Currently, reviewers write individual reviews and interdisciplinary team leader writes review
- Modern trial designs – call for statisticians to be deeply engaged at the inception rather than at the end
  - Master protocols
  - Bayesian methods
  - Need more cross-fertilization across disciplines
  - Engaged in stat reviews
- Medical divisions will be better placed to assist companies
  Example – she met with mom n=1 trial for antisense product.
  - Need to learn how to deal with drugs targeting only 10 patients world-wide, for example.
  - Need to find subsets of patients and target therapies, for example, in Type II diabetes. Different phenotypes.
  - Enough study subjects respond to get product to market. Want to understand subjects who don't respond.
- As we look at more data, can we put data in cloud and share.
  - Move transactional work with industry
  - Move data to cloud and standardize.
  - Use data instead of text as medium of exchange.
- Will be multi-year process. Need to move regulatory process to keep up with all the changes in drug development. She's very positive of direction of change. Will be profound shift within CDER and better ways to relate with external world

**<u>Q&A</u>**

**Q**: What is the role of statisticians in new paradigm.

**A**: Sentinel and FAERS require more intensive data analysis. Linking to registries, death certificates, medical records. Discussion of using AI to troll through everything. Will need statisticians to do this. EHR not high quality – better than what we have had before.

**Q:** Do you see statisticians helping with benefit-risk evaluations and including patient tradeoffs.
**A**: JW says she published papers on quantitative benefit-risk evaluations to show it can be done. Upsets clinical colleagues because you must translate into QALYs. But some think this is methodologically sound. Need to move to quantitative. Although not a hypothesis-based comparison, still valuable (e.g., responses for men are similar to responses for women). Eyeballing doesn't work. Need quantitative assessments to get people to focus on benefits and harms. The way we discuss today is inaccessible.

**Q:** With an issues-based approach, how do you know if you captured everything?
**A:** Not against each discipline having its own checklist. For example: clinpharm – did they get dose right.

---

## SESSION 1: PROMOTING COMPLEX INNOVATIVE TRIAL DESIGNS: AN UPDATE AND DISCUSSION ON THE PILOT MEETING PROGRAM

1. DIONNE PRICE, PHD, ACTING DEPUTY DIRECTOR, OFFICE OF BIOSTATISTICS, OTS, CDER FDA

Dr. Price *provided an update on the progress of the pilot meeting program* for complex innovative trial designs (CID).

Pilot Meeting Program Update for CID

- Launched in Aug 2018; pilot is 5 years.
- CDER/CBER selects 2 submissions per quarter
  - Sponsor will have 2 CID meetings
  - Sponsor agrees that FDA can use design as a case study for continuing ed and info sharing

FDA website for Complex Innovative Trial Designs Pilot Program:
https://www.fda.gov/drugs/development-resources/complex-innovative-trial-designs-pilot-program

Industry/public can submit questions online. FDA's answers are posted.

- Sponsor must have a pre-IND or IND
  - Intent to include CID in study
  - Provide substantial info
- Review timeline review: 240 days between first day of quarter and 2nd CID meeting
  - Day 45 – disclosure submission
- First pilot submission -- Wave Life Sciences. Publicly announced by Wave in press announcement
  - Duchenne multiple dystrophy
  - Placebo augmentation using a Bayesian modelling strategy
- Lessons learned
  - Ambitious timeline for both FDA and industry
  - Simulation plan and report needed for Meeting One Package
  - Need oversight steering committee

2. TELBA IRONY, PHD, DEPUTY DIRECTOR, OFFICE OF BIOSTATISTICS AND EPIDEMIOLOGY, CBER, FDA

To illustrate the role of simulations, Dr. Irony described an example where the proportional hazard ratio assumption was not met (e.g., delay in administration or delay in effectiveness).

Role of simulations:

- Simulations needed because closed formula for operating characteristics not often available for frequentist and Bayesian designs.
- Use of prior information – control group borrows strength from historical controls, external controls, natural history
    - Need priors – agree with FDA
    - Promising areas: peds, rare disease, small populations, safety, unmet medical need (use registries)
- Operating characteristics of interest:
    - Type I error
    - Power
    - Sample size, expected sample size
    - PTS
    - Prob of early success
    - Power
    - Etc.

3. VLADIMIR DRAGALIN, PHD, VICE PRESIDENT, HEAD OF QS CONSULTING, QUANTITATIVE SCIENCE JANSSEN R&D, AT JOHNSON & JOHNSON

    Title: "Industry perspective on the simulations practices for complex innovative designs"

Paper published in Biopharm Stats – includes industry and FDA co-authors.

Dr. Dragalin's presentation overlapped Dr. Irony's presentation: statistical properties and operating characteristics. I summarize here new points.

- Industry goals for a simulation report
    - Industry views a simulation report as a positive – report facilitates comparisons of pros and cons of candidate designs
    - Consistency and clarity of a clinical trial simulation report
    - Stimulates discussion and sharing
    - Increases awareness
- Simulation-guided clinical trial design
    - Highly iterative
    - Multiple design modifications, iterations with study team and multiple sets of scenarios, assumptions and metrics
    - Communication and clarity on the meaning of scenarios, values of nuisance parameters
    - Statistical validity, efficiency, financial measures
- The Biopharm Stats publication outlines simulation report content
- The simulation report is a living document
- He discussed reproducibility and validation of simulations

- Presented some recommendations from the ADSWG (adaptive designs study working group – DIA) simulation team
    - Details of simulations belong in simulation plan – not in the protocol in order to minimize operational bias.
    - Do not combine simulation report with SAP

4. PANEL DISCUSSION – Q&A:

Dr. Price:

- FDA has some flexibility for more than 2 meetings described in pilot program. So far, the 120 days between meetings seems reasonable. Not much time between when meeting granted and Meeting 1.
- No requirement to go through pilot program. Can go through traditional path.

Dr. Tom Louis stated simulations are needed for non-complex designs too. Need to be equally aggressive for designs we think we know about – but we don't.

Suggestions for convincing internal stakeholders are needed. Dr. Dragalin says a common team reaction is the process is too long.

The panelists discussed how to use CID more broadly for chronic conditions affecting many subjects.

- Dr. Price – a goal of program is to advance the use of CIDs in more therapeutic areas.
- FDA most interested in learning about the selected design. But good to know about designs that were considered but not pursued.

Software was also discussed. FDA plans to develop best practices for software. Potentially, standard software will be developed for designs that become more standard.

---

**SESSION 2: CHALLENGES AND OPPORTUNITIES FOR BAYESIAN METHODS IN CONFIRMATORY SETTING**

1. MARK ROTHMANN, PHD, ACTING DIRECTOR, DIVISION OF BIOMETRICS II, OFFICE OF BIOSTATISTICS, CDER, FDA

Dr. Rothmann discussed the following:

- How to communicate with non-statisticians
- Influence of a prior distribution
- Mixture priors
- Skeptical prior: data need to convince skeptic
- Power prior: refers to exponent, not probability of achieving a statistically significant result

He showed an example for pediatrics

2. MELANIE QUINTANA, PHD, STATISTICAL SCIENTIST, BERRY CONSULTANTS. LLC

- Berry Consultants see Bayesian designs for rare diseases, pediatrics, platform trials, others (stroke, peripheral arterial disease…)

- They developed a longitudinal model to predict final outcomes of a clinical study, using data external to the study. The idea was to use data outside trial, plus data from inside trial (different populations, non-contemporary randomized data – platform studies).

  Example: DIAN (dominantly inherited Alzheimer's)
  Rare <1%
  Early onset, 30 years old to 50 years old
  Opportunity to enroll preclinical patients who will certainly progress

3. KYLE WATHEN, PHD DIRECTORJANSSEN RESEARCH & DEVELOPMENT, LLC

Dr. Wathen discussed Bayesian designs for Hep C studies. These studies have consistency across historical controls, which makes Bayesian more plausible.

- Assumptions and simulations for initial scenarios
  - Historical studies
  - Non-inferiority margin
  - Standard of care (SOC) rates are high: .96 to .99. The larger the SOC, the larger the inflation of Type I error
- Simulations
  - Looked at different scenarios to see amount of Type I error rate
  - Presented a nice graphic that compared frequentist and Bayesian approaches. Contained color codes for comparisons.

Conclusions:
Bayesian design can bring innovation
Statisticians educated clinical team
Discussed design with HRAs
Pilot program likely to stimulate more discussion
More experience discussing the simulation report with the FDA is needed

4. TELBA IRONY, PHD, DEPUTY DIRECTOR, OFFICE OF BIOSTATISTICS AND EPIDEMIOLOGY, CBER FDA (DISCUSSANT)

Dr. Irony discussed each presentation then concluded with a summary.

Dr. Rothmann's presentation:
How to use adult results to extrapolate to pediatrics –
Pediatric decision tree is key
Challenges:
Subjectivity
How influential is the prior
Discounting the prior (several ways)
Arbitrary thresholds
Clinical input necessary: understanding
Need to explain to clinicians. Simulations important

Dr. Wathan's presentation (non-inferiority studies)
Used historical controls and natural history of disease to power a control group
Reduced trial duration and size
See Draft Guidance for Rare diseases – natural history study construction

Challenges
- Subjectivity
- How much to borrow
- Requires substantial clinical input
- Subjectivity more salient when results are borderline

Factors to support prior
- Large effect size of control
- Consistency across hx controls
- Lab tests
- RCTs with similar designs
- Large broad-based hx datasets
- Similar baseline characteristics

Dr. Quintana's presentation
Opportunities
Adaptive designs – insurance to account for deviations from assumptions

Dr. Irony's overall summary
Priors
- Subjective, requires agreement
- Come early to FDA
- Convince clinicians

Beware of arbitrary thresholds
Modeling – may provide lots of power if model assumptions hold true
Test assumptions: simulate, sensitivity analyses
Meet with regulators.

5. PANEL DISCUSSION:

- How to help clinicians?
  - Meet with clinicians to discuss their concerns and how priors work.
  - Show clinicians case studies and examples. More intuitive than p-value realm.
  - Bayesian resonates with clinicians because they like to pick a drug that has XX% of succeeding.
- Regarding options for discounting - what are the best ways to manage operating characteristics?
  - Can be done through decision analytics, patient utilities,
  - Consider utility approach. Compare designs to see which are preferable. Helpful to thinking, at the beginning when selecting design
- How to handle the situation where prior is much different from observed?
  - Perhaps the expert opinion wasn't good or historical control data didn't match.
  - Dr. Irony would not use prior that is not adaptive. Need to evaluate prior during interim looks. Should be prespecified in design along with solutions.

**SESSION 3: REAL WORLD DATA AND ITS FIT FOR REGULATORY PURPOSE**

Two speakers discussed "Pragmatic Trial Design and Execution". One was from industry, the other from FDA.

1. JAMES HARNETT, PHARMD, MS, SENIOR DIRECTOR, LEAD, REAL WORLD DATA AND ANALYTICS, PATIENT, AND HEALTH IMPACT, PFIZER INC

Dr. Harnett presented an industry perspective on the potential for randomized pragmatic clinical trials and the implications for label changes.

- Traditional RCTs are a powerful tool, but face challenges
  - <5% of eligible cancer patients enroll in RCTs
  - Physicians don't offer clinical trials to 50 to 80% of eligible patients
- Regulatory consideration of RWE beyond safety
  - How can we use clinical information to form decisions?
  - FDA active with partnerships – especially cancer. Example – avelumab. Used RWD to contexualize single arm trials
- How to generate RWE for regulatory decision making
  - Randomization, e.g., randomized pragmatic trials
  - Data source from clinical practices
- He presented 2 case studies, asking if RWD are fit for purpose
  - Jansen – schizophrenia. Has labeling claim.
  - GSK – Breo
- Will pragmatic clinical trials (PCTs) become standard?
  - Probably not
  - But will become more frequent with regulatory encouragement
  - A challenge: not typical investigators
- Barriers – see his slide
- Need blinded assessors
- Need to look at cost per data.
- Although RCTs will continue to be staple, with increasing availability of data and sources, there will be a variety of PCTs

2. DAVID MARTIN, MD, MPH, ASSOCIATE DIRECTOR FOR REAL WORLD EVIDENCE ANALYTICS, OMP, CDER, FDA

Dr. Martin discussed his views from his role as director for RWE at CDER/FDA.

- 21st Century Cures Act contains RWD/RWE requirements
  - Deadline was end of 2018 for producing a guidance document
    - How to use/evaluate RWE
    - Law says drugs but guidance extended to biologics
    - Substantial evidence
    - Comparative safety
  - Comments still open on RWE framework guidance
  - States substantial evidence standard not changed
- Potential uses of RWD/RWE in clinical studies – rare diseases, oncology
  - These typically use single arms – can use external controls to help.
  - Non-randomized, non-interventional studies

> FDA website - Real World Evidence:
> https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
>
> "This Website was designed to capture up-to-date information about the status of FDA activities around the development and use of RWD and RWE."
>
> Contains links to guidance document.

- Considerations in NDA/BLA submissions
  - Are RWD fit for use
  - Can trial/study design used to generate RWE provide adequate scientific evidence to answer or help
  - Public comments say data standards needed
- Discussed factors when considering a randomized clinical trial in RWD
- Example: "RELIANCE trial" – demonstration project – public/private partnerships. Want to pressure test assumptions. Working with PCORI to evaluate rofilumast vs azythromicin for prevention of COPD exacerbations.

3. FRANK W. ROCKHOLD, PHD, MSC, PROFESSOR OF BIOSTATISTICS, DUKE CLINICAL RESEARCH INSTITUTE, DUKE UNIVERSITY MEDICAL CENTER, UNITED STATES

Dr. Rockhold moderated the open discussion.

He says stats community needs to pay more attention to RWD, especially to address bias that most don't think about. Although E10 articulated these concepts, now we are applying them to a new world.

PCORI trials -- Whether randomized or not – all arms are compared to an active control. All contaminated and biased to null. When a difference occurs – do we believe it? If you don't find a difference, what do you do? Answer: We're still in the early days. Work needs to be done in advancing how we think about how we analyze data. There will be tremendous advancements. FDA is willing to accept less robust difference in end-of-life situations – benefit-risk. However, there's a tendency of a single standard for assessing evidence. FDA is behind the times. World of data analysis is rapidly changing. Machine learning and AI are leading to new insights. Reluctance within FDA to deal with non-inferiority studies in critical care, even in traditional RCT space. The RELIANCE trial is designed under PCORI.

A FLATIRON representative asked about the reproducibility of evidence and duplication of results of RWD studies
Answer (FDA representative): Been able to reproduce results with using similar entry criteria in a few studies (i.e., similar to RCTs). If we trust method, then maybe in the future we can say ok to RWD/RWE.
Answer (industry representative): Restrict data to what looks like RCT. Then if you have a slightly similar question, perhaps could have confidence in answer. Could use to supplement. HRT in post-menopausal women is always brought up as a counter example. Can't always define outcomes endpoint the same as in a RCT. How do we start opening the aperture to allow these data.

Dr. Rockhold: If I got different answer, could be different assumptions and answering different questions.

Dr. Tom Louis: If you have a high-quality observational study and results differ from what were expected or observed elsewhere (e.g., RCTs), ask why. For example, Women's Health Initiative study was re-analyzed and an answer was determined. When designing an observational study, collect data that will help with future needs and studies, even though the data may not be of interest at the time of the design.

---

**SESSION 4: ADVANCING PATIENT FOCUSED DRUG DEVELOPMENT**

1. PUJITA VAIDYA, MPH, SENIOR ADVISOR, PATIENT-FOCUSED DRUG DEVELOPMENT PROGRAM, OCD, CDER, FDA, UNITED STATES

Ms. Vaidya provided an update on the 4 guidance documents under development.

2. MARTIN HO, MS, ASSOCIATE DIRECTOR OF QUANTITATIVE INNOVATION, CDRH,FDA, UNITED STATES

Mr. Ho discussed use of patient preference information (PPI) to expand the labelling of certain technologies to include treatment at home without a care partner.

He gave case studies from CDRH and from CBER: clotting factors used in hemophilia and ongoing CBER patient input studies (Osteoarthritis – knees, sickle cell, and hard to control Type 1 diabetes).

3. ELIZABETH BUSH, MHS, DIRECTOR, PATIENT-FOCUSED OUTCOMES CENTER OF EXPERTISE, GLOBAL PATIENT OUTCOMES, ELI LILLY AND COMPANY, UNITED STATES

Ms. Bush discussed measurement strategy considerations.

She discussed tolerability as assessed by PROs vs clinician assessments.

Need to understand what makes sense to measure. For example, commercial decisions need to be based on good information, responder definitions, etc. Decisions are yes/no. However, if the product gets to market and it's unsuccessful, perhaps the company didn't measure the right thing or what was important.

Industry develops evidence packages for many stakeholders (e.g., PRO dossier for FDA, within company).

Industry doesn't have the same expectation of evidence for all stakeholders. For example, going into label vs what goes to payors. Principles are the same, but level of evidence differs.

Challenges in analyzing and interpreting PRO data

- QoL, PROs were handled in a separate section of protocol. Wasn't integrated into entire protocol
- Biostatisticians and psychometricians are not the same profession, nor do they have the same skills.

4. LAURA LEE JOHNSON, PHD, DIRECTOR DIVISION III, OFFICE OF BIOSTATISTICS, OTS, CDER, FDA, UNITED STATES

Dr. Johnson moderated a panel discussion and Q&A. She reminded the audience that FDA is working on a series of PFDD guidance documents. The PRO guidance document will be re-written.

Key comments on revised PRO guidance document:

- Guidance document should be a framework so FDA doesn't need to update guidance documents continuously. For example, don't need to explain IRT in detail
- Think of all stakeholders – e.g., patients, industry, academia. Guidance document needs to find balance.
- If it's a rule, industry will follow. Industry spends a lot of time reading tea leaves.

---

**SESSION 5: INNOVATIVE METHODS FOR ASSESSMENT OF COMPLEX GENERIC PRODUCTS**

1. STELLA C. GROSSER, PHD, DIVISION DIRECTOR, OFFICE OF BIOSTATISTICS, OTS, CDER, FDA, UNITED STATES

Title of presentation: The Bioequivalence of Complex Generic Drug Products

Dr. Grosser described the regulatory definition of a copy

- Modern system of generic drugs codified in 1984 Hatch-Waxman amendment – ANDA. A copy requires:
  - Need an approved reference product
  - Establish pharmaceutical equivalence
  - Establish bioequivalence
  - Adequately labelled and manufactured under GMP
- Therapeutic equivalent = pharmaceutical equivalent + bioequivalent

She described the type of evidence needed to establish bioequivalence. Many product-specific guidance documents available at FDA.gov. (https://www.accessdata.fda.gov/scripts/cder/psg/index.cfm)

Need to consider bioavailability at site of action.

GDUFA II introduced complex products.

Products with

- Complex routes of delivery
  - Topical dermatological products
  - Ophthalmic products
- Complex dosage forms
- Complex drug-device combinations

Complexity calls for additional statistical methods (beyond two one-sided tests)

She discussed clinical endpoint studies for complex generics

- Product-specific guidance documents recommend populations, design, endpoints, analyses
- Continuous endpoint
- Binary endpoints – e.g., difference in proportions should fall within +/- 20%

In some situations, complex generics call for non-inferiority. For example, epi pens. Want to compare error rates in use. Therefore, want generic to be non-inferior around error rate.

2. ROBERT A. LIONBERGER, PHD, DIRECTOR, OFFICE OF RESEARCH AND STANDARDS, OFFICE OF GENERIC DRUGS, CDER, FDA, UNITED STATES

Title: "Innovative approaches for complex generics"

The intention is to substitute a generic product for the brand product.

Pharmaceutical equivalents: same active ingredient, dosage form, strengths – typically no statistics needed for this
Bioequivalent: No significant differences in rate and extent of drug *at site of action.*

Complex generics defined in GDUFA II: large peptides, other dosage forms, all locally acting drugs (eye, nasal cavity, skin), drug-device combos with user interface considerations, abuse deterrent formulations.

When site of action is systemic, need PK studies

Delivered directly to sites of action – comparative endpoint BE studies, characterization-based approaches, weight of evidence (combined in vitro and in vivo)

Conclusion: opportunity to develop innovative BE methods for complex generics. Efficient decision making is essential to the success of the generic drug program.

---

**SESSION 6: MASTER PROTOCOLS: DESIGN CONSIDERATIONS AND OPERATIONAL EXAMPLES (CASE-STUDIES)**

1. SCOTT M BERRY, PHD, PRESIDENT AND SENIOR STATISTICAL SCIENTIST, BERRY CONSULTANTS LLC, UNITED STATES

Dr. Berry referenced Woodcock and LaVange article on Master Protocols:
https://www.nejm.org/doi/full/10.1056/NEJMra1510062

He discussed I-SPY 2

2. SOFIA PAUL, PHD, SENIOR DIRECTOR, GSK, UNITED STATES

Title: Master Protocols: Design Considerations and Operational Examples (Case-Studies)

Dr. Paul discussed a platform trial designed to determine a P2 dose for multiple myeloma treatment.

3. JONATHAN HADDAD, MPH, DIRECTOR, BIOSTATISTICS, GLAXOSMITHKLINE, UNITED STATES

Dr. Haddad moderated a panel discussion (2 panelists)
Here are some key points:

- Dr. Berry re-emphasized that FDA encourages innovation, citing the CID pilot program as an example.
- He pointed out that some platform studies may have many placebo arms, but not a common control, with the exception of oncology which tends to have a common control.
- In response to a question on how to handle changes over time in SOC, Dr. Berry said this happened in I-SPY 2. For breast cancer, they introduced a new SOC.
- You can use patients from, say, 5 or 6 years ago to help inform estimate of efficacy despite change to SOC. Estimate relative efficacy across arms by modelling drift over time. Can take advantage of the history of all enrolled patients and overlap of multiple arms. Model network to estimate each individual drug.

---

**SESSION 7: CANCER PATIENT JOURNEYS AND ESTIMANDS**

1. JOHN SCOTT, PHD, DIRECTOR, DIVISION OF BIOSTATISTICS, OBE, CBER, FDA, UNITED STATES

Title: "Cancer Patient Journeys and Estimands

Dr. Scott discussed four elements of an estimand: population, variable, population-level summary and a discussion of intercurrent events. He emphasized how to provide information to patients that is meaningful to them.

Population:
A patient asks, is this trial about people like me?
Oncology studies have many possibilities for defining a population

- Tumor type
- Tumor location
- Stage/severity
- Treatment resistance
- Demographics

Variable:
Am I getting information about experiences I care about?
Oncology studies include endpoints for survival, progression, response, other clinical variables, QoL
Although we care about survival for evaluating treatment, not clear if patients care most about survival
If endpoint is a surrogate – what does that mean to patient: reasonably likely to predict clinical outcome

Intercurrent events:
What do results mean?
Every TA has own intercurrent events challenges
Oncology (example):

- Progression
- Treatment crossover
- Toxicity
- Randomized and never treated

Population summary
What are the results?
Summary should provide a basis for making treatment comparisons – should be clinically interpretable parameters
Sensible it should be aligned with test: can we align summary with test stat?
For proportional hazards
    Do hazard ratios (HR) meet this criterion? Think how to explain HR to patient
    Do median differences? If median short, maybe what I care about is the tail of the survival distribution
For non-proportional hazards
    HRs undefined
    Medians don't fully capture differences

Patient-friendly population summaries

- To understand what happened in the trial, I want multiple percentiles, with estimates of precision
- To understand what's going to happen to me, I want: probabilities, the probability that a patient like me will live 3 months, 6 months, etc.

2. SATRAJIT ROYCHOUDHURY, PHD, SENIOR DIRECTOR, BIOSTATISTICS, PFIZER INC, UNITED STATES

**Title:** Insights from a Cross-Pharma Working Group on Estimands in Oncology Trials

Dr. Roychoudhury referenced a PSI-authored manuscript:
https://onlinelibrary.wiley.com/doi/abs/10.1002/pst.1745

She reviewed 5 estimand strategies

- Treatment policy
- Hypothetical
- Composite or transformed
- Principal strata
- On treatment

She described potential Intercurrent events in oncology studies:

- Administrative censoring
- Early study withdrawal
- Post-progression treatment cross-over
- Dose reduction
- On-treatment withdrawals

See her slides for an example of estimands in post-progression treatment cross-over

| How to handle Intercurrent event | Estimand |
|---|---|
| Ignore | treatment policy |
| Associate with progression | composite policy |
| Censored | hypothetical |
| More complex (IPW) | hypothetical |

There is an oncology estimands working group

- Started 2018

- 31 members, 19 companies
- Centered in Europe
- Key goals and activities of 5 sub teams:
    - Causal estimands in time-to-event studies
    - Treatment switching
    - Censoring mechanisms and their impact on interpretation of estimands
    - Case studies in solid tumors
    - Case studies in hematology

3. KAPILDEB SEN, SENIOR DIRECTOR BIOSTATISTICS, NOVARTIS PHARMA, UNITED STATES

Title: Patient Journeys and Estimands: A case study of CAR-T in Lymphoma

A living drug designed to target CD19. CAR-T cells are manufactured from patient's cells

For their single arm P2 study, the challenge was how to handle patient drop-outs because of the situation where a patient drops out before CAR-T infusion. Here are some possible scenarios:

- Patient's condition deteriorated
- CAR-T manufacturing failed.
- Patient received chemo and improved when CAR-T was available.

What is the treatment effect of interest?
The strategy?
Only those who took CAR-T
Bridging chemo followed by CAR-T infusion treatment strategy

What is the proper baseline?
Timing of baseline
Evidence of disease at baseline?

EMA – focused on enrolled patients with evidence of disease at enrollment.
FDA – focused on infused patients with evidence of disease prior to infusion

<u>P3 study design:</u>
CAR-T treatment not available at baseline
Complex treatment strategies

He showed a very informative table with 3 columns headed intercurrent event, handling strategy, justification

4. PANEL DISCUSSION

The panel was presented this question: Is the agency expecting to see estimands in protocols? Especially in oncology and other therapeutic areas?

Dr. Temple suggested a reread of ICH E9, which describes what you are to do with endpoint, dropouts, reasons. He asked, what is the difference between E9 and the amendment, when E9 already says what you should do. De. Scott says there is no difference in behavior. R(1) emphasizes to think ahead. CBER seen a small number of INDs that have estimands. Only at Step 2 of document. Implementation details will come out with final version.

Dr. Scott says the Step 2 document discusses time-to-event. However, no clear guidance on non-inferiority. Dr. Temple says Step 2 says not to do NI studies.

Dr. Temple says typical oncology study is time to progression (TTP). Lots of treatments showed effect on TTP but didn't show time to survival (OS) because of crossovers. What is different now? Difference between test drug and control will disappear because subjects who progress can take effective drug.

The panel discussed how far back in drug development to go to estimands. Phase 1 is a bit early! Clinical publications should reflect estimands more often. Will take a while for people to become familiar with terms.

Dr. Temple discussed how to present individual results. Our guidance talks about CDFs, but not many do so. What about tails. Are there estimands that can help? Impressive to have 4 patients, for example, to have no disease after 4 years.

Dr. Bill Wang commented on the need to bring together safety estimands and efficacy estimands.

---

**SESSION 8: SENIOR LEADERS TOWN HALL: BRIDGING STATISTICAL SCIENCE WITH DATA SCIENCE IN DRUG DEVELOPMENT**

Speakers:
ALOKA CHAKRAVARTY, PHD, ACTING DIRECTOR, OFFICE OF BIOSTATISTICS, OFFICE OF TRANSLATIONAL SCIENCES, CDER, FDA, UNITED STATES

ERIC GIBSON, PHD, VP, GLOBAL HEAD BIOSTATISTICAL SCIENCES AND PHARMACOMETRICS, NOVARTIS PHARMACEUTICALS CORPORATION, UNITED STATES

CYRUS HOSEYNI, PHD, VP & GLOBAL HEAD OF STATISTICS AND DECISION SCIENCES, JANSSEN R&D, UNITED STATES

LISA LUPINACCI, ASSOCIATE VICE PRESIDENT, MERCK AND CO., INC., UNITED STATES

JOHN SCOTT, PHD, DIRECTOR, DIVISION OF BIOSTATISTICS, OBE, CBER, FDA, UNITED STATES

RAM TIWARI, PHD, DIRECTOR, DIVISION OF BIOSTATISTICS, CDRH, FDA, UNITED STATES

Q1: Big Data vs Statistics
Statisticians focus on questions need to be answered and develop research plan and analyses
Statisticians need to focus on Big Data

- Numerous analyses of secondary use of data
- Need better understanding of how data generated
- Need to focus more on design aspects, where the data come from
    - Example: an analysis done of studies conducted in Europe

Big role for senior software engineers – can mentor statisticians. For example, documentation, communication, understand each other's code. Follow same best practices regarding

Q2: What are the best ways for statisticians to engage with other types of data scientists to get their expertise, understanding and truly add value

Statisticians understand methods well, but don't necessarily understand black-box of machine learning and AI. New FDA statisticians have better understanding of AI and ML.

Different skills are needed: statisticians, computer skills, secondary use of data, need to meet ahead of experiment, design challenges. Rather than interpreting post-hoc exploratory analyses

Advise grad programs on classes: nearest neighbors, cluster analyses, training courses for the future

Statisticians need to become more comfortable about research direction vs research questions. For example, digitized MRI database. Maintain an open-mind with computer scientists, machine learners. Learn the language in order to be able to speak openly/easily with other team members.

One company has many groups of data scientists for different type of company-wide data such as, for example, performance reviews and manufacturing. Opportunity for statisticians to collaborate with groups across company at an enterprise level.

In a decade, pharma companies will be data science companies (e.g., Novartis). Much data being generated. Statisticians are in decline. However, statisticians can lead and bring people together. We have the confidence to lead.

Statisticians need to be open-minded and navigate attitudes – for example, pre-specified exploratory analyses vs more open. More inter-disciplinary teamwork. Use right tools for the job.

---

**SESSION 9: UNHEALTHY SAFETY ASSESSMENT: MOVING TOWARDS BETTER CHARACTERIZATION OF PATIENT HARMS**

1. PETER P. STEIN, MD, DIRECTOR, OFFICE OF NEW DRUGS, CDER, FDA, UNITED STATES

Title: Improved Characterization of Patient Harms: A Regulatory Perspective

He reviewed FDA safety assessment 'steps'
Statisticians help identify subgroups at risk for safety signals
Consider implications for post-approval use
Determine appropriate management
Reviewed statistical challenges for safety assessments
- P3 studies not typically powered for safety
- Pooling

Need early involvement during IND development: design

2. CHRISTOPHER B. GRANGER, MD, PROFESSOR OF MEDICINE; DIRECTOR, CARDIAC INTENSIVE CARE UNIT, DUKE UNIVERSITY MEDICAL CENTER, UNITED STATES

Title: Unhealthy Safety Assessment: Moving Towards Better Characterization of Patient Harm

3. STEFFEN UNKEL, DEPARTMENT OF MEDICAL STATISTICS, UNIVERSITY MEDICAL CENTER GÖTTINGEN, GERMANY

Title: Estimands and the Analysis of Adverse Events in the Presence of Varying Follow-Up Times Within the Benefit Assessment of Therapies

Unfortunately, I had to step out of the session and didn't hear these 2 speakers.

**SESSION 10: OPPORTUNITIES FOR STATISTICAL LEADERSHIP AND INNOVATIONS IN RARE DISEASE THERAPIES**

1. LUCAS KEMPF, MD, ACTING ASSOCIATE DIRECTOR, RARE DISEASES PROGRAM, OND, CDER,FDA, UNITED STATES

Title: Statistical Challenges in Trial Design in Rare Diseases: It is more than just a numbers game

Registries are typically used to help design studies – not for historical controls
SMART trial designs require early statistical input
They require good natural history data
Endpoint: what is meaningful and what is possible: Patient input – "Nothing about us without us"

2. KELLEY KIDWELL, ASSOCIATE PROFESSOR OF BIOSTATISTICS, UNIVERSITY OF MICHIGAN, SCHOOL OF PUBLIC HEALTH, UNITED STATES

Title: SMART Design for Rare Disease Clinical Trial Research

Primary goal: compare treatments by pooling
Statistics in Medicine paper last year discussed SMART studies and compared Bayesian and frequentist models: https://onlinelibrary.wiley.com/doi/10.1002/sim.7900
Used Bayesian joint statistical model (BJSM)
Compared treatment with highest posterior mean versus treatment with second highest posterior mean
Will give presentation at SCT meeting

3. PANEL DISCUSSION

MARSHALL SUMMAR, MD, CHIEF, DIVISION OF GENETICS AND METABOLISM, CHILDREN'S NATIONAL MEDICAL CENTER, UNITED STATES

KAREN LYNN PRICE, PHD, MA, SENIOR RESEARCH ADVISOR, STATISTICAL INNOVATION CENTER ELI LILLY AND COMPANY, UNITED STATES

LAURA LEE JOHNSON, PHD, DIRECTOR DIVISION III, OFFICE OF BIOSTATISTICS, OTS, CDER, FDA, UNITED STATES

Discussion around how to put information in labeling, including RWE/RWD, different sources of data.

Need a system of knowledge management among different groups developing natural history studies and registries for a particular indication. Need to share best practices and develop standards. It's important to take extra measures to reduce variability within small studies – for example, take multiple blood pressure readings instead of 1.

Networking events among health care providers and families can improve patient outcomes in studies. Some studies, like for adenovirus, are n=1 and done. What is the implication for future studies of more effective drugs/treatments. As early as possible, sit down with HCPs and understand practicalities of disease. May find better markers of disease and outcomes. If families buy-in, would get better cooperation and higher quality studies.

---

**SESSION 11: CAUSALITY, ARTIFICIAL INTELLIGENCE AND BIG DATA**
ELIAS BAREINBOIM, PROFESSOR, PURDUE UNIVERSITY, UNITED STATES
Title: Causal Inference and Data-Fusion

Dr. Bareinboim recommended the following book, which reviews causal inference over the last 40 years:
**THE BOOK OF WHY: THE NEW SCIENCE OF CAUSE AND EFFECT HARDCOVER – MAY 15, 2018, BY JUDEA PEARL AND DANA MACKENZIE**

He presented a nice discussion on graphical views of causal inference with an example of observational data (effect of exercise on cholesterol). Reminded audience, a transparent representation of the mechanisms underlying the phenomenon under investigation is needed.

Mark van der Laan
He presented targeted machine learning for generating real-world evidence from observational data and gave a roadmap of statistical learning

Targeted learning is more than just machine learning. Targeted learning is needed in order to get appropriate statistical inference.

Causal modeling helps get estimate. Machine learning and targeting minimize the statistical gap.

Super learning: Everyone has a different model, for example, to predict survival. Which of the many options is best to use. What is the solution – going to use all of it. He recommends building library of the algorithms. Then let data figure out which is best. Set aside 10% of data, and train. Choose the winner. Can weight the algorithms.