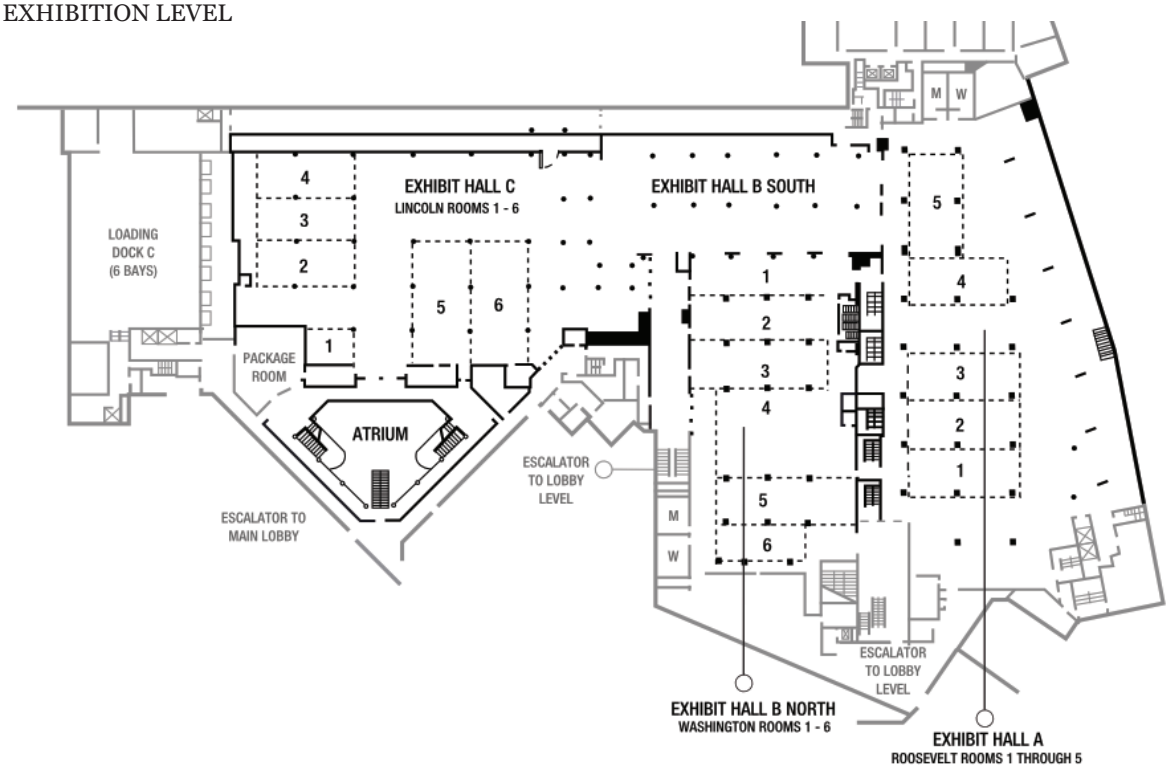
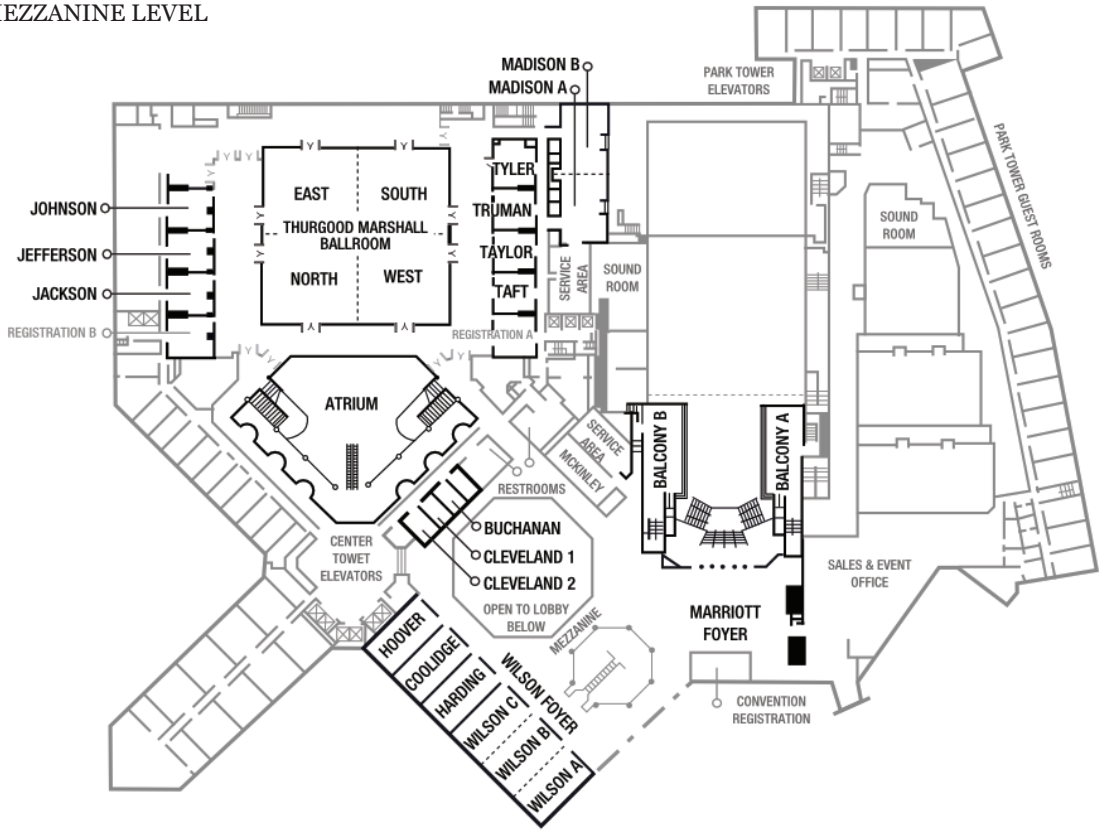
2015 ASA Biopharmaceutical Section Statistics Workshop

September 16–18, 2015
Marriott Wardman Park Hotel
Washington, DC

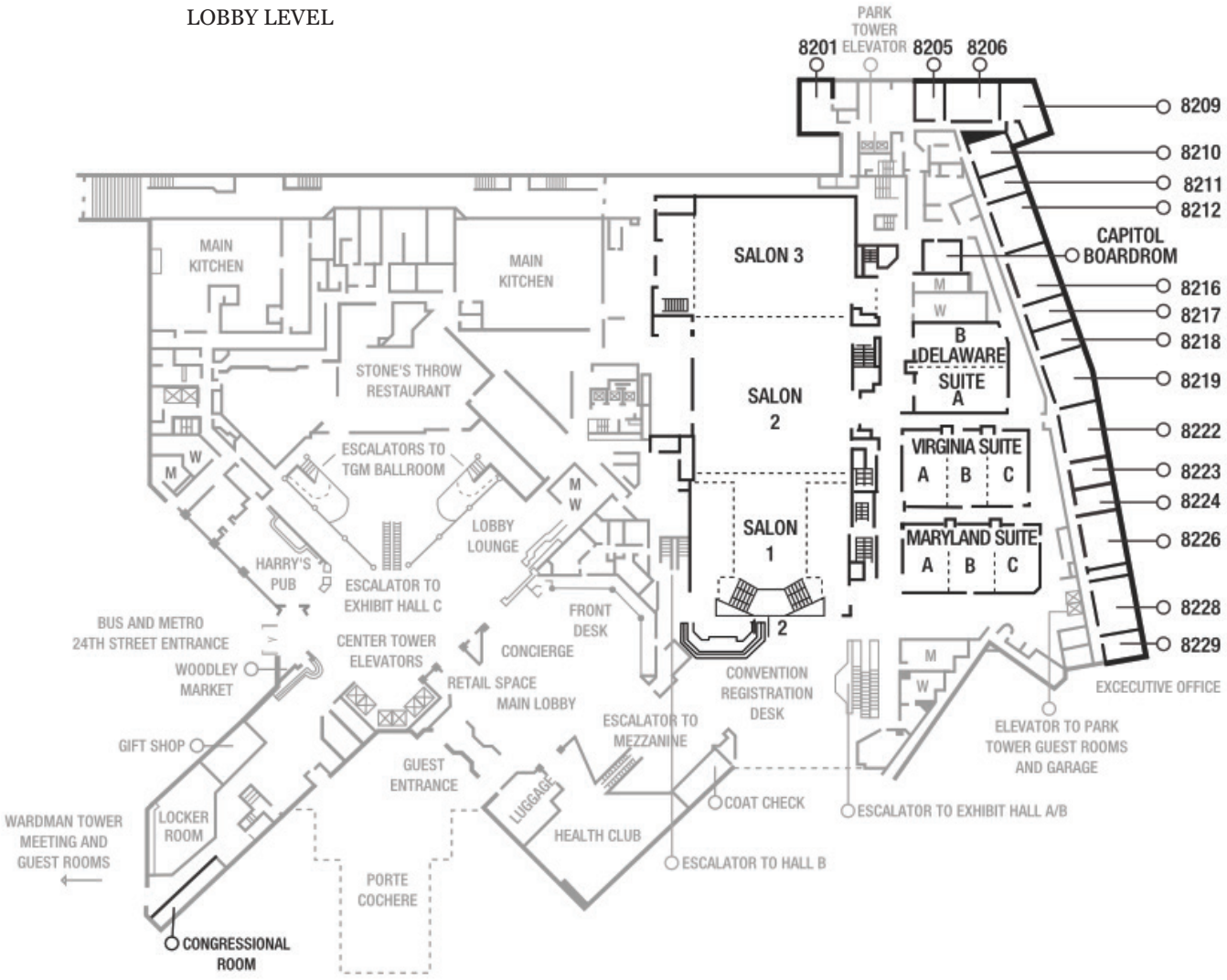
EXHIBITION LEVEL



MEZZANINE LEVEL



LOBBY LEVEL


PARK TOWER ELEVATOR
8201
8205
8206
8209
8210
8211
8212
CAPITOL BOARDROM
8216
8217
8218
8219
8222
8223
8224
8226
8228
8229
MAIN KITCHEN
MAIN KITCHEN
SALON 3
SALON 2
SALON 1
STONE'S THROW RESTAURANT
ESCALATORS TO TGM BALLROOM
LOBBY LOUNGE
HARRY'S PUB
ESCALATOR TO EXHIBIT HALL C
FRONT DESK
CONCIERGE
CENTER TOWER ELEVATORS
RETAIL SPACE MAIN LOBBY
BUS AND METRO 24TH STREET ENTRANCE
WOODLEY MARKET
GIFT SHOP
GUEST ENTRANCE
ESCALATOR TO MEZZANINE
CONVENTION REGISTRATION DESK
DELAWARE SUITE A
VIRGINIA SUITE A B C
MARYLAND SUITE A B C
EXCECUTIVE OFFICE
ELEVATOR TO PARK TOWER GUEST ROOMS AND GARAGE
WARDMAN TOWER MEETING AND GUEST ROOMS
LOCKER ROOM
LUGGAGE
HEALTH CLUB
COAT CHECK
ESCALATOR TO EXHIBIT HALL A/B
ESCALATOR TO HALL B
PORTE COCHERE
CONGRESSIONAL ROOM

DEAR COLLEAGUE,

On behalf of the ASA Biopharmaceutical Section, welcome to the 2015 ASA Biopharmaceutical Section Statistics Workshop and to the nation's capital, Washington DC. We look forward to an exciting three-day workshop, starting with eight half-day short courses on Wednesday and followed by two days of talks in plenary, parallel, and town hall formats.

In his 2015 State of the Union address, President Barack Obama announced a precision medicine initiative with a goal *"to enable a new era of medicine through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized treatments."* To this end, the workshop begins with plenary presentations and a panel discussion about the future of precision medicine, and this theme persists in many of the sessions and presentations throughout the workshop.

In addition, the program offers a wide range of relevant topics for statisticians working on or interested in medical product development and the related regulatory environments. These topics include:

- Bayesian methods
- Benefit-risk
- Biomarkers
- Cardiovascular safety in type II diabetes
- CMC applications
- Data standards and transparency
- Estimands and sensitivity analysis for missing data
- Generics and biosimilars
- Incorporating patient perspectives
- Medical devices and diagnostics
- Patient enrichment
- Statistical leadership
- Subgroups
- Methodologies for vaccine and veterinary development

Building on a practice that began at the 2014 workshop, organizers, presenters, or panelists from any session had the opportunity to submit an article related to their session for consideration in an upcoming issue of *Statistics in Biopharmaceutical Research (SBR)*.

Furthermore, two *SBR* invited sessions—Incorporating Patient Perspectives in the Medical Product Life Cycle (PS6d) and Concerns with Reanalysis for Ongoing Data Transparency Initiatives (PS3a)—will contribute articles. This special issue will be dedicated entirely to the 2015 workshop.

Roundtable lunches will take place Thursday. Luncheon discussion topics are divided into zones and housed in different rooms.

Please refer to the program to find the location of your topic discussion. Your roundtable assignment should be included with your badge packet. Come ready to participate!

Finally, we want to extend our sincere gratitude to our steering committee members, organizing committee members, consultants, session chairs, session organizers, speakers, panelists, discussants, and workshop volunteers for their dedication and diligence in creating a truly outstanding program. In addition, we want to thank our ASA meeting planner, Christina Link, and the ASA staff for their tremendous efforts in coordinating the workshop logistics. This conference could not happen without their contributions of time, talent, dedication, and insight.

We look forward to meeting you and hope you have a rewarding, informative, and enjoyable time participating in this workshop!

Sincerely,
Wei Zhang and Richard C. Zink
Co-chairs

2015 ASA Biopharmaceutical Section Statistics Workshop

Workshop Co-chairs
Wei Zhang (FDA/CVM)
Richard Zink (SAS)

Session Proposal Subcommittee
Yu Cheng (University of Pittsburgh)
Terri Johnson (FDA/CDRH)
Caiyan Li (Takeda)
Jingyu Luan (FDA/CDER)
Ed Luo (PTC Therapeutics)
Theodore Lystig (Medtronic)
Erik Pulkstenis (MedImmune)
Virginia Recta (FDA/CVM)
John Scott (FDA/CBER)

Short Course Subcommittee
Chris Holland (Amgen)
Telba Irony (FDA/CDRH)
Bret Musser (Merck)
Mark Levenson (FDA/CDER)
Camille Orman (Janssen)
Yue Shentu (Merck)
Yuqing Tang (FDA/CDRH)

Roundtable Luncheon Subcommittee
Freda Cooner (FDA/CDER)
Rakhi Kilaru (PPD)
Inna Perevozskaya (Pfizer)

Mixer Subcommittee
Chris Holland (Amgen)
Anastasia Ivanova (UNC Chapel Hill)
Hugh Rand (FDA/CFSAN)

CMC Subcommittee
Stan Altan (JNJ)
J. David Christopher (Merck)
Xiaoyu Dong (FDA/CDER)
Meiyu Shen (FDA/CDER)
Yi Tsong (FDA/CDER)
Harry Yang (MedImmune)

Follow @AmstatNews and tag #ASABiopharm

Steering Committee
FDA
Freda Cooner (FDA/CDER)
Telba Irony (FDA/CDRH)
Terri Johnson (FDA/CDRH)
Shiowjen Lee (FDA/CBER)
Mark Levenson (FDA/CDER)
Jingyu Luan (FDA/CDER)
Hugh Rand (FDA/CFSAN)
Virginia Recta (FDA/CVM)
John Scott (FDA/CBER)
Yuqing Tang (FDA/CDRH)

Industry
Chris Holland (Amgen)
Rakhi Kilaru (PPD)
Caiyan Li (Takeda Global Research & Develop. Center, Inc.)
Ed Luo (PTC Therapeutics)
Theodore Lystig (Amgen)
Cristiana Mayer (JNJ)
Bret Musser (Merck)
Camille Orman (JNJ)
Inna Perevozskaya (Pfizer)
Erik Pulkstenis (MedImmune)
Yue Shentu (Merck)

Academia
Yu Cheng (University of Pittsburgh)
Anastasia Ivanova (UNC Chapel Hill)

Consultants
Greg Campbell (FDA/CDRH)
James Chen (FDA/NCTR)
Lisa LaVange (FDA/CDER)
Anna Nevius (FDA/CVM)
Estelle Russek-Cohen (FDA/CBER)

ASA Staff
Heather Alexander, Customer Service Representative
Christina Link, Meetings Planner
Kathleen Wert, Director of Meetings

Room Assignments

TL1, TL2, TL3, TL4, TL5, TL6, TL7, TL8	MARRIOTT BALCONY B
TL9, TL10	TRUMAN
TL11, TL12	TYLER
TL13, TL14, TL15, TL16	TAFT
TL17, TL19, TL20	WILSON A
TL21, TL22, TL23, TL24, TL25, TL26, TL27, TL28	MARRIOTT BALCONY A
TL29, TL30, TL31, TL32	WILSON B
TL33, TL34, TL35	MCKINLEY
TL36, TL37, TL38, TL39, TL40	MADISON A
TL41, TL42, TL43, TL44, TL45, TL46	MADISON B
TL47, TL48	TAYLOR
TL49, TL50, TL51	WILSON C
LUNCH ONLY (NO DISCUSSIONS)	LINCOLN 2,3,4

Wednesday, September 16

7:30 a.m. – 5:30 p.m.
REGISTRATION
Registration Area B

8:30 a.m. – 12:00 p.m.
MORNING SHORT COURSES

Short Course 1: An Overview of Statistical Considerations in Personalized Medicine: Concept and Methodology
Instructor(s): *Meijuan Li, FDA*
Thurgood Marshall North East

Short Course 2: Handling Missing Data in Clinical Trials
Instructor(s): *Sonia Davis, The University of North Carolina at Chapel Hill; Michael O'Kelly, Quintiles*
Thurgood Marshall South West

Short Course 3: Equivalence and Similarity Testing
Instructor(s): *Shein-Chung Chow, Duke University; Yi Tsong, FDA/CDER*
Madison AB

Short Course 4: Introduction to PK/PD Modeling for Statisticians
Instructor(s): *Yaming Hang, Biogen Idec; Alan Hartford, AbbVie*
Wilson ABC

9:45 a.m. – 10:00 a.m.
REFRESHMENT BREAK
Thurgood Marshall Foyers

12:00 p.m. – 1:30 p.m.
LUNCH ON OWN

1:30 p.m. – 5:00 p.m.
AFTERNOON SHORT COURSES

Short Course 5: Dose-Finding in Drug Development: Methods and Implementation, with Focus on MCP-Mod
Instructor(s): *Frank Bretz, Novartis; José C. Pinheiro, Johnson & Johnson*
Thurgood Marshall North East

Short Course 6: Statistical Strategies for Clinical Development of Personalized Medicines
Instructor(s): *Cong Chen, Merck*
Thurgood Marshall South West

Short Course 7: Bayesian Adaptive Phase I Oncology Trials: Methodology and Implementation
Instructor(s): *Beat Neuenschwander, Novartis Pharma AG; Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals*
Madison AB

Short Course 8: Designing Observational Comparative Studies Using Propensity Score Methodology in Regulatory Settings
Instructor(s): *Donald Rubin, Harvard University; Lilly Q Yue, FDA/CDRH*
Wilson ABC

3:30 p.m. – 3:45 p.m.
REFRESHMENT BREAK
Thurgood Marshall Foyers

Thursday, September 17

7:00 a.m. – 5:30 p.m.
REGISTRATION
Registration Area B

7:00 a.m. – 7:45 a.m.
CONTINENTAL BREAKFAST
Thurgood Marshall Foyers

8:00 a.m. – 8:15 a.m.
OPENING REMARKS FROM WORKSHOP CO-CHAIRS

8:15 a.m. – 9:45 a.m.
PLENARY SESSION
Thurgood Marshall Ballroom

Plenary Session 1 – The Future of Precision Medicine
Organizer(s): *Wei Zhang, FDA/CVM; Richard C. Zink, JMP Life Sciences, SAS Institute*

Precision Medicine Initiatives at FDA
Lisa LaVange, FDA

Micro-Randomized Trials and mHealth
Susan Murphy, University of Michigan

9:45 a.m. – 10:00 a.m.
REFRESHMENT BREAK
Thurgood Marshall Foyers

10:00 a.m. – 11:30 a.m.
PLENARY SESSION
Thurgood Marshall Foyers

Plenary Session 2 – Panel Discussion on the Future of Precision Medicine
Organizer(s): *Wei Zhang, FDA/CVM; Richard C. Zink, JMP Life Sciences, SAS Institute*
Panelists: *Greg Campbell, formerly of FDA/CDRH; Cong Chen, Merck; Lisa LaVange, FDA; Susan Murphy, University of Michigan; Estelle Russek-Cohen, CBER FDA; Richard Simon, National Cancer Institute*

11:45 a.m. – 1:00 p.m.
ROUNDTABLE LUNCHEON (TICKETED EVENT)

1:15 p.m. – 2:30 p.m.
PARALLEL SESSIONS
Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

2:45 p.m. – 4:00 p.m.
PARALLEL SESSIONS
Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

4:00 p.m. – 4:15 p.m.
REFRESHMENT BREAK
Thurgood Marshall Foyers

4:15 p.m. – 5:30 p.m.
PARALLEL SESSIONS

Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

5:45 p.m. – 6:45 p.m.
MIXER

Thurgood Marshall Foyers

Friday, September 18

7:30 a.m. – 1:00 p.m.
REGISTRATION

Registration Area B

7:30 a.m. – 8:15 a.m.
CONTINENTAL BREAKFAST

Thurgood Marshall Foyers

8:30 a.m. – 9:45 a.m.
PARALLEL SESSIONS

Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

10:00 a.m. – 11:15 a.m.
PARALLEL SESSIONS

Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

11:15 a.m. – 12:45 p.m.
LUNCH ON OWN

12:45 p.m. – 2:00 p.m.
PARALLEL SESSIONS

Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

2:15 p.m. – 3:30 p.m.
PARALLEL SESSIONS

Thurgood Marshall North
Thurgood Marshall South
Thurgood Marshall East
Thurgood Marshall West
Lincoln 5
Lincoln 6

Adaptive Design

TL1 Logistics and Implementation of Adaptive Trial Designs — *Eva Miller, inVentiv Health Clinical*

TL2 Enriching Patient Population by Response: Placebo Run-In, Randomized Withdrawal, Sequential Parallel Comparison Design (SPCD), and Twice-Enriched Design (TED) — *Anastasia Ivanova, The University of North Carolina at Chapel Hill*

TL3 Key Characteristics in Bayesian Adaptive Design — *Xin Fang, FDA/CDRH*

TL4 Challenges and Opportunities for Statisticians in Planning/Implementing Adaptive Trial Designs — *Nan Shao, Covance, Inc.*

TL5 H0 *P*-Value–Based Futility Decision and Other Seemingly Inappropriate Methods — *Stan Lin, FDA/CBER*

TL6 Adaptive Designs in Unblinded Studies? — *Jie (Jack) Zhou, FDA/CDRH*

TL7 Are Statisticians Ready to Implement Increasing Number of Platform Trials? — *Emelita de Leon-Wong, PPDI*

TL8 Adaptive design in Medical Device Trials — *Peter Lam, Boston Scientific*

Bioequivalence, Generics, and Biosimilars

TL9 Statistical Issues and Methods in Biosimilar — *Jin Xu, Merck*

Biomarkers

TL10 Precision Medicine: Statistical Issues, Trial Design, and Regulatory Aspects — *Amir Handzel, AstraZeneca*

Comparative Effectiveness

TL11 Relevance and Data Accessibility for Network Meta-Analyses for Comparative Effectiveness Research Using Patient-Level Randomized Clinical Trial Data — *Leiya Han, PPD*

TL12 Statistical Assessment of Comparative Effectiveness in Clinical Trials — *Isaac Nuamah, Janssen R&D*

Diagnostics

TL13 Precision Studies for In-Vivo Devices — *Bipasa Biswas, FDA*

DSMB/InterimAnalysis/ Advisory Committee

TL14 Incorporating Futility into a Phase 3 Outcomes Trial Governed by a Data Monitoring Committee — *Richard Davies, GSK*

Early Phase Trials

TL15 Robust Decision-Making in Early-Stage Clinical Development — *Yanli Zhao, MedImmune/ AstraZeneca; Erik Pulkstenis, MedImmune*

High-Dimensional Data (e.g., Pharmacogenomics)

TL16 Best Practices in Next-Generation Sequencing Methodology with Impact on High-Dimensional Findings — *Justin Davis, AbbVie*

Medical Devices

TL17 How to Treat Site in Clinical Trials: Fixed or Random? — Chul Ahn, FDA-CDRH

TL19 Making Sense of Sensors — *Vadim Zipunnikov, The Johns Hopkins University*

Meta-Analysis

TL20 Bayesian Meta-Analysis and Meta-Analysis for Stroke and Myeloma — *Xiaoping Liu, Zhongnan Hospital of Wuhan University*

Missing Data

TL21 The Prevention and Treatment of Missing Data in Clinical trials: How Far Have We Come? — *Gosford Sawyerr, Janssen Pharmaceuticals*

TL22 Practical Issues with MMRM — *Dalong Huang, Takeda Development Center Americas, Inc.*

TL23 Impact of Missing Data and Their I mputations in Long-Term Treatment of Chronic Auto-Immune Diseases — *Achim Guettner, Novartis Pharam AG*

TL24 Missing Data Analysis Planning in Late-State Clinical Trials: A Check-Up on Current Practices — *Davis Gates, Merck*

TL25 Investigating Product Complaints: Pitfalls of Working with Manufacturing Data — *Aaron Spence, BIOVIA*

Modeling and Simulation

TL26 Validation of Predictive Modeling in Observational Studies — *Rui Li, Quintiles; Zahouhui Su, Quintiles*

TL27 The Interface Between Statistical and PKPD Modeling and Simulation — *Matthew Rotelli, Eli Lilly and Company*

TL28 Analyses of Longitudinal Clinical Data with Time-Varying Covariates — Rong Liu, Eli Lilly and Com*pany; Qianyi Zhang, Eli Lilly and Compan***y**

Noninferiority

TL29 Bayesian and Frequentist Approaches to Noninferiority Clinical Trials — *Carl DiCasoli, Bayer Healthcare Pharmaceuticals*

TL30 Noninferiority Trial with Survival Endpoints — *Mengdie Yuan, FDA; Elena Rantou, FDA/CDER*

Observational Studies

TL31 The Impact of EU Post-Approval Safety Surveillance Studies (PASS) — *Charles Liss, AstraZeneca Pharmaceuticals*

TL32 Statistical Considerations for Handling Treatment Switches in Observational Studies — *William Hawkes, Quintiles RWLPR*

Oncology

TL33 PFS: Central vs. Local — *Lihui Zhao, Novartis*

TL34 Noninferiority in Cancer Trials — *Tingting Yi, Novartis*

TL35 Challenges and Opportunities of Statistics in Oncology Immunotherapy — *Yi He, Celldex Therapeutics*

Other

TL36 Statistical Intellectual Property — *Philip Lavin, Lavin Consulting LLC*

TL37 A Comprehensive Review of the Multiple-Sample Tests for the Three General Data Types — *Ying Yang, FDA/CDRH*

TL38 Leadership and Career Development for Junior Statisticians Working on Clinical Trials — *Lei Gao, Sanofi*

TL39 Crossover Design in Clinical Studies — *Tao Wang, Eli Lilly and Company*

TL40 We the People of the Biopharm Section … Chat with the Chair-Elect — *B. Christine Clark, QDS*

Patient-Reported Outcomes and Patient Preferences

TL41 Patient-Reported Outcomes in Oncology — *Laura Fernandes, FDA*

TL42 Incorporating Patient Preferences Evidence into Regulatory Considerations — *Martin Ho, FDA/CDRH*

TL43 PRO and COA Experiences: Regulatory and Patient Priorities and Processes — *Laura Johnson, FDA*

Regulatory Topics/Guidances

TL44 Pathway for Antibiotics: Revisiting Endpoints and Designs — *Prasanna Ambati, PPD*

TL45 Next-Generation Sequencing Diagnostic Tests — *Peggy Wong, Merck*

TL46 Challenges and Good Practices to Improve the Quality of Therapeutic Device Submissions — *Manuela Buzoianu, FDA/CDRH*

Role of Statisticians

TL47 Promotion of Involvement of Statistician and Statistical Analysis of Risk-Based Monitoring in Clinical Trials — *Xiaoqiang Xue, Quintiles*

Safety

TL48 Blinded and Unblinded Evaluation of Aggregate Safety Data During Clinical Development — *Bill Wang, Merck*

Therapeutic Area Specific Topic

TL49 Risk Stratification Strategies to Identify Low-Risk Patients in Cardiovascular Clinical Trials — *Juliana Ianus, Janssen R&D; CV Damaraju, Janssen R&D*

TL50 Suicidal Ideation and Behavior: Design and Analysis of Clinical Trials — *Pilar Lim, Janssen R&D; Rosanne Lane, Janssen R&D*

Quality/Validation

TL51 IWRS: Interactive Web Response System: Looking Beyond Randomization and Medication Kit Assignment — *Rama Melkote, Janssen, R&D ; Kim Cooper, Janssen, R&D*

Thursday, September 17
1:15 p.m. – 2:30 p.m.

PS1A

Statistical Experiences on Subgroup-Stratified, Biomarker-Stratified, or Enrichment Trials

Thurgood Marshall North

Organizer(s): *Eva Miller, inVentiv Health Clinical; Liming Dong, FDA; Yuanjia Wang, Columbia University; Xiting (Cindy) Yang, FDA/CDRH*
Chair(s): Xiting (Cindy) Yang, FDA/CDRH
Speakers: *Bo Huang, Pfizer Inc.; Guoxing (Greg) Soon, FDA/CDER/OTS/OB; Richard Simon, National Cancer Institute*

PS1B

Current and Future Role of the Clinical Statistician in the World of Data Transparency

Thurgood Marshall South

Organizer(s): *Jeffrey Joseph, Theorem Clinical Research; Stephen Wilson, FDA/CDER/OTS/OB/DBIII; Vivian Shih, AstraZeneca; Yueqin Zhao, FDA*
Chair(s): *Stephen Wilson, FDA/CDER/OTS/OB/DBIII*
Speakers: *Michael Pencina, Duke Clinical Research Institute; Marc Buyse, IDDI Inc.; Jeffrey Gardner, Janssen R&D*
Discussant(s): *Stephen Wilson, FDA/CDER/OTS/OB/DBIII*

PS1C

Big Data/Big Analytics: Challenges and Opportunities in Pre- and Post-Market Medical Product Evaluations Utilizing National/International Registries

Thurgood Marshall East

Organizer(s): *Yunling Xu, FDA/CDRH; Nelson Lu, FDA/CDRH; Charles Darby, Statistical Consultant; Rakhi Kilaru*
Panelists: *Jesse Berlin, J&J; Chunrong Cheng, CBER/FDA; Rima Izem, OB/CDER/FDA; Ted Lystig, Medtronic Inc.; Bram Zuckerman, CDRH/FDA*
Speakers: *Donald Rubin, Harvard University; Lilly Yue, FDA/CDRH; Nelson Lu, FDA/CDRH; Yunling Xu, FDA/CDRH*

PS1D

Analytical Similarity: Current Statistical Issues in Biosimilar Product Development

Thurgood Marshall West

Organizer(s): *Meiyu Shen, FDA; Harry Yang, MedImmune LLC*
Chair(s): *Harry Yang, MedImmune LLC*
Discussant(s): *Yi Tsong, FDA/CDER; Rick Burdick, Amgen Inc.*

PS1E

DMCs and Adaptive Clinical Trials: Considerations in Balancing Safety and Trial Integrity

Lincoln 5

Organizer(s): *Michelle Detry, Berry Consultants; Jie (Jack) Zhou, FDA/CDRH; Greg Ball, AbbVie; Zhuang Miao, FDA*
Chair(s): *Michelle Detry, Berry Consultants*
Speakers: *Roger Lewis, Berry Consultants; Greg Campbell, Consultant; Thomas Cook, University of Wisconsin*

PS1F

Advanced Multiple Testing Methodologies for Confirmatory Trials

Lincoln 6

Organizer(s): *Xuan Liu, AbbVie; Freda Cooner, FDA/CDER; Anthony Rodgers, Merck; Julia Jingyu Luan, FDA/CDER*
Chair(s): *Xuan Liu, AbbVie*
Speakers: *Walter Offen, AbbVie; Frank Bretz, Novartis; Willi Maurer, Novartis; Xiaolei Xun, Novartis; Mohammad Huque, FDA/CDER*
Discussant(s): *Hsien-Ming James Hung, FDA*

Thursday, September 17
2:45 p.m. – 4:00 p.m.

PS2A

Logical Inference on Treatment Efficacy in Subgroups and Their Combinations in Personalized Medicine Development

Thurgood Marshall North

Organizer(s): *Ying Ding, University of Pittsburgh; Xiang Ling, FDA/CDER; Jason Hsu, Eli Lilly and Company/Ohio State University; Thomas Birkner, FDA*
Chair(s): *Jason Hsu, Eli Lilly and Company/Ohio State University*
Speakers: *Yi Liu, Takeda Pharmaceuticals; Ying Ding, University of Pittsburgh*
Discussant(s): *Hsien-Ming James Hung, FDA*

PS2B

Subgroup Analysis Under Rising Regulatory Emphasis: Fundamentals and Challenges

Thurgood Marshall South

Organizer(s): *Yijie Zhou, AbbVie; Bo Yang, AbbVie; Weiya Zhang, FDA; Yifan Wang, FDA*
Speakers: *Lu Cui, AbbVie; Shufang Liu, AbbVie; Janet Wittes, Statistics Collaborative*
Discussant(s): *Bob Temple, FDA*

PS2C

Large Trials for Major Adverse Cardiovascular Events

Thurgood Marshall East

Organizer(s): *Aloka Chakravarty, FDA/CDER; Bret Musser, Merck; Olga Marchenko, Quintiles*
Chair(s): *Richard Zink, JMP Life Sciences, SAS Institute*
Speaker(s): *Olga Marchenko, Quintiles*
Panelists: *Aloka Chakravarty, FDA/CDER; Qi Jiang, Amgen; José Pinheiro, J&J; Estelle Russek-Cohen, FDA/CBER*

PS2D

Quality and Quality Metrics

Thurgood Marshall West

Organizer(s): *Stan Altan, J&J; Liang Zhao, FDA*
Speakers: *Lawrence Yu, FDA/CDER/OPQ; Helen Strickland, GSK*

PS2E

Current Statistical Issues in Biosimilar Product Development

Lincoln 5

Organizer(s): *Bo Jin, Pfizer Biotechnology Clinical Development; Sungwoo Choi, FDA/CDER; Jason Liao, Novartis; Xin Gao, FDA/CDER*
Chair(s): *Joshua Chen, Sanofi Pasteur; Eric Chi, Amgen Inc.*
Speakers: *Bo Jin, Pfizer Biotechnology Clinical Development; Kerry Barker, Pfizer Inc.; Cassie Dong, FDA; Gregory Levin, FDA*

PS2F

Adaptive Enrichment Design: A Way to Achieve the Goal of Personalized Medicine?

Lincoln 6

Organizer(s): *Min (Annie) Lin, FDA/CBER; Zhiwei Zhang, FDA/CDRH; Feng Liu, GSK; Inna Perevozskaya, Pfizer Inc.*
Chair(s): *Meijuan Li, FDA; Feng Liu, GSK*
Speaker(s): *Cong Chen, Merck; Nicole Li, Merck; Zhiwei Zhang, FDA/CDRH; Michael Rosenblum, Johns Hopkins Bloomberg School of Public Health*

Thursday, September 17
4:15 p.m. – 5:30 p.m.

PS3A

Concerns with Reanalysis for Ongoing Data Transparency Initiatives

Thurgood Marshall North

Organizer(s): *Theodore Lystig, Medtronic, Inc.*
Chair(s): *Michael Hale, Amgen*
Speaker(s): *Jonathan Hartzel, Merck; Estelle Russek-Cohen, FDA/CBER; Sara Hughes, GSK*

PS3B

Bayesian Subgroup Analysis: Opportunities and Challenges in Unmet Medical Need

Thurgood Marshall South

Organizer(s): *Margaret Gamalo, FDA/CDER; David Ohlssen, Novartis; Helen Zhou; Freda Cooner, FDA/CDER*
Speaker(s): *Satrajit Roychoudhury, Novartis; Santosh Sutradhar, Novartis; Kert Viele, Berry Consultants; Liz Krachey, Berry Consultants; Gene Pennello, FDA/CDRH*
Discussant(s): *Ravi Varadhan, The Johns Hopkins Center on Aging and Health*

PS3C

Town Hall Session: Roles of Statisticians in Academia, Regulatory, and Pharmaceuticals Industry

Thurgood Marshall East

Organizer(s): *Yulan Li, Novartis; Guoxing (Greg) Soon, FDA/CDER/OTS/OB; Yanming Yin, FDA/CDER/OTS/OB; Keaven Anderson, Merck*
Chair(s): *Yulan Li, Novartis; Guoxing (Greg) Soon, FDA/CDER/OTS/OB*
Speaker(s): *Janet Wittes, Statistics Collaborative; Robert Califf, FDA; Lisa LaVange, FDA; Jeffrey Helterbrand, Roche*
Panelists: *Robert Califf, FDA; Lisa LaVange, FDA; Jeffrey Helterbrand, Roche; Janet Wittes, Statistics Collaborative; Bob Temple, FDA; Tom Fleming, University of Washington; Sonia Davis, University of North Carolina at Chapel Hill; Ramachandran Suresh, GSK*

PS3D

Continuing Discussion: Statistical Considerations for Continuous Manufacturing Processes

Thurgood Marshall West

Organizer(s): *David Christopher, Merck; Yi Tsong, FDA/CDER; Helen Strickland, GSK*

PS3E

New Statistical Methods for Risk Assessment

Lincoln 5

Organizer(s): *Zhiwei Zhang, FDA/CDRH; Mei-Ling Lee, University of Maryland; Mengdie Yuan, FDA; Greg Ball, AbbVie*
Chair(s): *Zhiwei Zhang, FDA/CDRH*
Speaker(s): *Yueqin Zhao, FDA; Lei Shen, Eli Lilly and Company; Mei-Ling Lee, University of Maryland*

PS3F

Practical Experiences Using Meta-Analysis

Lincoln 6

Organizer(s): *Anna Nevius, FDA/CVM; Steven Radecki, ASA; Ed Luo; Lisa Rodriguez, FDA/CVM*
Chair(s): *Virginia Recta, FDA/CVM*
Speaker(s): *Laura Hungerford, FDA/CVM; Emily Smith, FDA/CVM; Steven Radecki, ASA; Junshan Qiu, FDA/CDER; Anna Nevius, FDA/CVM*

Friday, September 18
8:30 a.m. – 9:45 a.m.

PS4A

Advancing Personalized Medicine Using Innovative Subgroup Identification Methods

Thurgood Marshall North

Organizer(s): *Qi Tang, AbbVie; Rong (Rachel) Chu, Agensys, Inc.; Yun Wang, FDA/CDER/OTS/OB/DB5; Anna Sun, FDA*
Chair(s): *Qi Tang, AbbVie*
Speaker(s): *James Chen, FDA/NCTR; Wei-Yin Loh, University of Wisconsin-Madison; Xin Huang, AbbVie*
Discussant(s): *Lei Shen, Eli Lilly and Company*

PS4B

Modeling and Simulation at the FDA and in Industry: Collaboration Among Statisticians, Modelers and Pharmacometricians

Thurgood Marshall South

Organizer(s): *Cristiana Mayer, J&J; Shiowjen Lee, FDA; Alan Hartford, AbbVie; Misook Park, FDA*
Chair(s): *Cristiana Mayer, J&J*
Speaker(s): *Tarek Haddad , Medtronic; Rajesh Nair, FDA/CDRH; Laura Thompson, FDA; Telba Irony, FDA; Adam Himes, Medtronic Inc.; José Pinheiro, J&J; Chyi-Hung Hsu, Janssen R&D; Vikram Sinha, FDA*

PS4C

Messy Data Issues in Evaluation of Bioequivalence

Thurgood Marshall East

Organizer(s): *Wanjie Sun, FDA; Stella Grosser, FDA; Mark Shiyao Liu, Mylan Inc.; Charles DiLiberti, Montclair Bioequivalence Services, LLC*
Chair(s): *Julia Jingyu Luan, FDA/CDER*
Speaker(s): *Wanjie Sun, FDA; Aotian Yang, GWU; Stella Grosser, FDA; Carol Kim, FDA; Charles Bon, Biostudy Solutions, LLC; Lindsey Katz, Biostudy Solutions, LLC*
Panelists: *Lisa LaVange, FDA; Yi Tsong, FDA; Shein Chong Chow, Duke; Pina D'Angelo, Noum Pharmaceutical Research Services*

PS4D

Recent Innovations in the Development and Application of Statistical Designs for Early-Phase Oncology Trials

Thurgood Marshall West

Organizer(s): *Yuan Ji, NorthShore University HealthSystem/The University of Chicago; Adam Hamm, Theorem Clinical Research, Inc.; Hui Zhang, FDA; Xian Zhou*
Chair(s)*: Sue-Jane Wang, FDA*
Speaker(s): *Peter Thall, MD Anderson Cancer Center; Inna Perevozskaya, Pfizer Inc.; Linda Sun, Merck; Christine Gause, Merck*
Discussant(s): *Lindsay Renfro, Mayo Clinic; Lei Nie, FDA/CDER*

PS4E

Meeting the Ebola Challenge

Lincoln 5

Organizer(s): *Estelle Russek-Cohen, FDA/CBER; Deepak Khatry, MedImmune; Dionne Price, FDA/CDER; James Lymp, Genentech*
Chair(s): *Deepak Khatry, MedImmune*
Speaker(s): *Ivan Chan, Merck; Kenneth Liu, Merck; Lori Dodd, NIAID; Michael Proschan, NIAID*
Discussant(s): *Estelle Russek-Cohen, FDA/CBER; Dionne Price, FDA/CDER*

PS4F

Minimizing Bias in Medical Device Trials Through Study Design and Data Analysis

Lincoln 6

Organizer(s): *Laura Lu, FDA/CDRH; Theodore Lystig, Medtronic, Inc.; Chia-Wen Ko, FDA*
Chair(s): *Laura Lu, FDA/CDRH*
Speaker(s): *Yunling Xu, FDA/CDRH; Heng Li, FDA/CDRH/OSB; Vandana Mukhi, FDA/CDRH/OSB; Nelson Lu, FDA/CDRH; Lilly Yue, FDA/CDRH; Peter Lam, Boston Scientific; Thomas Love, Case Western Reserve University*

Friday, September 18
10:00 a.m. – 11:15 a.m.

PS5A

Recent Developments and Considerations for Personalized Medicine: Follow-On 'Me Too' Companion Diagnostic Devices

Thurgood Marshall North

Organizer(s): *Yuying Jin, FDA/CDRH; Laura Yee, FDA/CDRH; Alicia Toledano; Kuang-Lin He, Fujirebio Diagnostics, Inc.*
Chair(s): *Qin Li, FDA/CDRH*
Speaker(s): *Xiao-Hua Zhou, University of Washington; Yunqi Bu, University of Washington; James Ranger-Moore, Roche Tissue Diagnostics; Crystal Schemp, Roche Tissue Diagnostics; Meijuan Li, FDA*

PS5B

Innovative Designs and Advanced Statistical Methodologies for Rare Disease Clinical Trials

Thurgood Marshall South

Organizer(s): *Yeh-Fong Chen, FDA; Hope Knuckles, Abbott; Laura Johnson, FDA; Jeffrey Krischer, University of South Florida*
Chair(s): *Yeh-Fong Chen, FDA*
Speaker(s): *Roy Tamura, University of South Florida; Min Min, FDA; GQ Cai, GSK*

PS5C

Role of Statisticians in CDISC Data Standards from Developers to Users

Thurgood Marshall East

Organizer(s): *Weiya Zhang, FDA; Deborah Bauer, Sanofi; Stephen Wilson, FDA/CDER/OTS/OB/DBIII; Peter Mesenbrink, Novartis*
Chair(s): *Deborah Bauer, Sanofi*
Speaker(s): *Peter Mesenbrink, Novartis; Weiya Zhang, FDA*
Panelists: *Stephen Wilson, FDA/CDER; Susan Kenny, Maximum Likelihood, Inc.; Chris Holland, Amgen; Deborah Bauer, Sanofi*

PS5D

ICH E9 R1 Defining the Estimand and Sensitivity Analysis

Thurgood Marshall West

Organizer(s): *Estelle Russek-Cohen, FDA/CBER; Brent Burger, PAREXEL International; Guoying Sun, FDA*
Chair(s): *Joan Buenconsejo, AstraZeneca*
Speaker(s): *Devan Mehrotra, Merck; Thomas Permutt, FDA*

PS5E

Using Historical Data in Clinical Trials: Synthesis of Truth with Uncertainty

Lincoln 5

Organizer(s): *Kerri Schoedel, Alteros Research Partners, Inc.; Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals; Pandurang Kulkarni, Eli Lilly and Company; Margaret Gamalo, FDA/CDER*
Chair(s): *Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals*
Speaker(s): *Beat Neuenschwander, Novartis; Jason Connor, Berry Consultants; Kristine Broglio, Berry Consultants; Manuela Buzoianu, FDA/CDRH*
Discussant(s): *Sujit Ghosh, North Carolina State University*

PS5F

Analysis and Interpretation of Human Abuse Potential Study Data

Lincoln 6

Organizer(s): *Yahui Hsueh, FDA; Hope Knuckles, Abbott; Kerri Schoedel; Chen Ling, FDA*
Chair(s): *Wei Liu, FDA/CDER/OTS/OB/DBVI; Marta Sokolowska, Grunenthal USA*
Speaker(s): *Chen Ling, FDA; Naama Levy-Cooperman, Altreos Research Partners; Kerri Schoedel; Susan Spruill, Applied Statistics and Consulting*

Friday, September 18
12:45 p.m. – 2:00 p.m.

PS6A

Statistical Considerations of Delayed Treatment Effects in Cancer Vaccine Trials

Thurgood Marshall North

Organizer(s): *Shenghui Tang, FDA; Zhenzhen Xu, FDA/CBER; Marc Buyse, IDDI Inc.; Jonathan Norton, MedImmune*
Chair(s): *Shenghui Tang, FDA*
Speaker(s): *Zhenzhen Xu, FDA/CBER; Jianliang Zhang, MedImmune; Erik Pulkstenis, MedImmune; Daowen Zhang, North Carolina State University*

PS6B

Designing Bioequivalence Studies for the Evaluation of Generic Drugs: Addressing Challenges Arising from Different Sources of Variability

Thurgood Marshall South

Organizer(s): *Elena Rantou, FDA/CDER; Fairouz Makhlouf, FDA/CDER; CV Damaraju, Janssen R&D; Susan Huyck, Merck*
Chair(s): *Fairouz Makhlouf, FDA/CDER*
Speaker(s): *Charles DiLiberti, Montclair Bioequivalence Services, LLC; Elena Rantou, FDA/CDER; Shein-Chung Chow, Duke University*

PS6C

Summarizing Case Studies to Learn and Improve Confirmatory Adaptive Trial Design and Implementation

Thurgood Marshall East

Organizer(s): *Weili He, Merck; Paul Gallo, Novartis; Xuefeng Li, FDA/CDRH; Min (Annie) Lin, FDA/CBER*
Chair(s): *Paul Gallo, Novartis; Xuefeng Li, FDA/CDRH*
Speaker(s): *Eva Miller, inVentiv Health Clinical; Weili He, Merck; Paul Gallo, Novartis*
Panelists: *Sue-Jane Wang, FDA; Greg Campbell, formerly of FDA/CDRH; Boguang Zhen, FDA/CBER; Jerry Schindler, Merck; Eva Miller, inVentiv Health Clinical; Weili He, Merck*

PS6D

Incorporating Patient Perspectives in the Medical Product Life Cycle

Thurgood Marshall West

Organizer(s): *Bennett Levitan, Janssen R&D; Scott Braithwaite, New York University; Telba Irony, FDA; Martin Ho, FDA/CDRH*
Speaker(s): *Telba Irony, FDA; Bennett Levitan, Janssen R&D; Scott Braithwaite, New York University*

PS6E

Common Statistical Issues FDA Encounters

Lincoln 5

Organizer(s): *Heng Li, FDA/CDRH/OSB; Vandana Mukhi, FDA/CDRH/OSB; Brent Burger, PAREXEL International; Adam Hamm, Theorem Clinical Research, Inc.*
Chair(s): *Heng Li, FDA/CDRH/OSB*
Speaker(s): *Shiowjen Lee, FDA; Laura L. Fernandes, FDA/CDER/OTS/OB/DBV; Xu Yan, FDA/CDRH; Heng Li, FDA/CDRH/OSB; Vandana Mukhi, FDA/CDRH/OSB*

PS6F

Missing Data in Diagnostic Device Studies: Methods and Case Studies

Lincoln 6

Organizer(s): *Bipasa Biswas, FDA; Xuan Ye, FDA; Vicki Petrides, Abbott; Kristen Meier, Illumina, Inc.*
Chair(s): *Bipasa Biswas, FDA*
Speaker(s): *Xiao-Hua Zhou, University of Washington; Yuqing Tang, FDA; Hope Knuckles, Abbott*
Discussant(s): *Gene Pennello, FDA/CDRH*

Friday, September 18
2:15 p.m. – 3:30 p.m.

PS7A

Platform Trials and Master Protocols: New Adaptive Designs Advancing Personalized Medicine

Thurgood Marshall North

Organizer(s): *Ohad Amit, GSK; Cristiana Mayer, J&J; Rajeshwari Shridhara, FDA; Lei Nie, FDA/CDER*
Chair(s): *Teri Ashton, GSK*
Speaker(s): *Lijun Zhang, FDA; Shenghui Tang, FDA; J. Kyle Wathen, Janssen R&D; Scott Berry, Berry Consultants*

PS7B

Poolable or Non-Poolable: Challenges and Solutions

Thurgood Marshall South

Organizer(s): *Ying Yang, FDA/CDRH; Yu Zhao, FDA/CDRH; Minjung Yoon; Ying Yan, Helsinn*
Chair(s): *Yunling Xu, FDA/CDRH*
Speaker(s): *Shun Zhenming, Sanofi-Aventis; Yu Zhao, FDA/CDRH; Ying Yang, FDA/CDRH; Joe Massaro, Boston University*

PS7C

Emerging Topics in Benefit-Risk Assessment

Thurgood Marshall East

Organizer(s): *Weili He, Merck; Qi Jiang, Amgen; John Scott, FDA/CBER; Xuefeng Li, FDA/CDRH*
Chair(s): *Weili He, Merck; Xuefeng Li, FDA/CDRH*
Speaker(s): *Qi Jiang, Amgen; Weili He, Merck; John Scott, FDA/CBER*
Panelists: *John Scott, FDA/CBER; Telba Irony, FDA; Ellis Unger, FDA/CDER; Qi Jiang, Amgen*

PS7D

Bayesian Assessment of Benefit-Risk Balance in Drug Development

Thurgood Marshall West

Organizer(s): *Maria Costa, GSK; Yueqin Zhao, FDA; Carl DiCasoli, Bayer Healthcare Pharmaceuticals; Min Min, FDA*
Chair(s): *Yueqin Zhao, FDA*
Speaker(s): *Deborah Ashby, Imperial College London; Ram Tiwari, FDA; Ian Hirsch, AstraZeneca*

PS7E

Statistical Considerations in Evaluating Imaging-Based Devices

Lincoln 5

Organizer(s): *Jincao Wu, FDA/CDRH; Jingjing Ye, FDA; Alicia Toledano; Jeffrey Joseph, Theorem Clinical Research*
Chair(s): *Jincao Wu, FDA/CDRH*
Speaker(s): *Nancy Obuchowski, Cleveland Clinic Foundation; Lucy McGowan, Vanderbilt University; Jennifer Bullen, Cleveland Clinic Foundation; Yuying Jin, FDA/CDRH; Meijuan Li, FDA; Qin Li, FDA/CDRH*

PS7F

Use of Phase 2 Interim Analysis to Expedite Drug Development Decisions

Lincoln 6

Organizer(s): *Jenny Huang, Genentech; Qi Xia, Genentech; Qin Li, FDA/CDRH; Norberto Pantoja-Galicia, FDA*
Chair(s): Jenny Huang, Genentech; Qi Xia, Genentech
Speaker(s)*: Qi Xia, Genentech; Yuan Shen, FDA/CDER/OTS/OB; Raji Sridhara, FDA/CDER/OTS/OB*
Discussant(s): *Daniel Sargent, Mayo Clinic*

SHORT COURSES

Wednesday, September 16
8:30 a.m. – 12:00 p.m.

Short Course 1: An Overview of Statistical Considerations in Personalized Medicine: Concept and Methodology

Thurgood Marshall North East

Organizer(s): *Yuqing Tang, FDA*
Instructor(s): *Meijuan Li, FDA*

The term "personalized medicine" is often described as providing "the right patient with the right drug at the right dose at the right time." More broadly, "personalized medicine" may be thought of as the tailoring of medical treatment to the individual characteristics, needs, and preferences of a patient during all stages of care, including prevention, diagnosis, treatment, and follow-up. This short course will provide a general overview of the concept and statistical methodology related to personalized medicine. The course will begin with a general discussion of statistical issues related to personalized medicine with the second part of the course focusing on concepts and principals for evaluating and planning companion diagnostic device studies.

A key component of personalized medicine is companion diagnostics that measure biomarkers (e.g., protein expression, gene amplification, or specific mutations). For example, most of the recent attention concerning molecular cancer diagnostics has been focused on the biomarkers of response to therapy, such as KRAS mutations in metastatic colorectal cancer, EGFR mutations in advanced Non-small cell lung cancer, and BRAF mutations in metastatic malignant melanoma. The presence or absence of these markers is directly linked to the response rates of particular targeted therapies with small-molecule kinase inhibitors or antibodies. Therefore, testing and evaluating for these markers has become a critical step in the target therapy of the above-mentioned tumors. For companion diagnostics devices, we will discuss device's indications for use, study designs, performance measures, and statistical methods of data analysis for both clinical and analytical validation studies of companion diagnostic devices. Real case examples will also be discussed.

Short Course 2: Handling Missing Data in Clinical Trials

Thurgood Marshall South West

Organizer(s): *Richard Zink, JMP Life Sciences, SAS Institute*
Instructor(s): *Sonia Davis, The University of North Carolina at Chapel Hill; Michael O'Kelly, Quintiles*

This half-day course looks at missing data in clinical trials. The course is based on *Clinical Trials with Missing Data: A Guide for Practitioners*, a recently published book to which both trainers contributed. The course starts with an illustrative example of clinical data related to Parkinson's disease. Basic concepts pertaining to missing data are explained. The assumption of missing at random (MAR) is described and shown in action for this data. The idea of missing not at random (MNAR) is also introduced. Regulatory thinking on missing data is summarized. This leads to the three main pillars of the course. The first pillar is a full introduction to the direct likelihood approach known as mixed models for repeated measures (MMRM). Examples of SAS code are presented that implement MMRM for the same Parkinson's disease data. The MMRM approach is used to implement the MAR assumption. The second pillar of the course covers multiple imputation. The key ideas of multiple imputation are described. The workings of multiple imputation are illustrated with SAS code. This multiple imputation methodology will be applied to data from a clinical trial in Major Depressive Disorder that is downloadable from *www.missingdata.org.uk*. Like MMRM, multiple imputation can be used to implement the MAR assumption; but multiple imputation is very flexible and can implement a wide variety of assumptions about missing data. The third pillar of the course shows how, by controlling the steps of multiple imputation, the statistician can in a very simple way implement rather sophisticated assumptions about missing data. One type of assumption is described in further detail—the assumption that post-withdrawal outcomes from the experimental arm may be modeled by observations from the control arm. Attendees will gain from this course a thorough knowledge of issues pertaining to missing data in clinical trials. Attendees will also gain an understanding of a variety of statistical approaches for handling missing data, and how to implement those approaches in practice.

Short Course 3: Equivalence and Similarity Testing

Madison AB

Organizer(s): *Yi Tsong, FDA/CDER*
Instructor(s): *Shein-Chung Chow, Duke University; Yi Tsong, FDA/CDER*

The objective of an equivalence study is to demonstrate if the two drug products are equivalent. However, "equivalence" may be defined in various ways. It may mean that the two drugs lead to similar responses; or to similar efficacies (over placebo) within a defined margin of equivalence. Depending on which of the two objectives, margin determination, study design, measurement and test may be different. In a study with only teat and reference products or treatments, the equivalence measure may be difference in means or ratio of means. Equivalence of two products/treatments is then demonstrated by showing that the measurement of difference

is between the pre-specified lower and upper margins of equivalence. Therefore, the null hypothesis of interest is that the difference measurement is either lower than the lower margin or larger than the upper margin. In another word, one may demonstrate equivalence by showing that the difference is laying between the two margins. The complete test may be performed with two sets of hypotheses. In a three products/treatment (placebo, test and reference) study, in addition to the equivalence test, the assay sensitivity of reference is established by showing reference is superior to placebo; efficacy of test is demonstrated by showing test is superior to placebo. Equivalence test is used in in-vivo studies for equivalence assessment of a generic product or a post-market change of a new product. It is also used in in-vitro studies for equivalence assessment of chemical or physical parameters. It is also used in clinical trials for equivalence assessment of patient treatment responses of the two products.

In this presentation, we will give an overview of the equivalence tests in terms of in vivo, in vitro or therapeutically equivalence; bioequivalence and biosimilarity.

Short Course 4: Introduction to PK/PD Modeling for Statisticians

Wilson ABC

Organizer(s): *Yaming Hang, Biogen Idec; Alan Hartford, AbbVie*
Instructor(s): *Yaming Hang, Biogen Idec; Alan Hartford, AbbVie*

Pharmacokinetic/pharmacodynamic (PK/PD) modeling using nonlinear mixed-effects models, also commonly called pharmacometrics, has been performed for many years. It is important for decision making in drug development and has impact on drug labels. This modeling is usually performed by scientists from a variety of backgrounds with very different levels of statistical training. However, statisticians should play an equal and important role in PK/PD modeling; there is a broad range of statistical issues in this field that can benefit from statistician's input. With wider acceptance of model-based drug development, statisticians need to at least be able to review modeling and simulation plans, results, and inferences to ensure correct implementation of statistical methodology and to appreciate the value of such analyses in the drug development process. In an effort to make this field more accessible to statisticians, this short course introduces concepts and methods for using nonlinear mixed-effects models for examining relationships between PK and PD endpoints while bridging the differences in terminology.

Wednesday, September 16
1:30 p.m. – 5:00 p.m.

Short Course 5: Dose-Finding in Drug Development: Methods and Implementation, with Focus on MCP-Mod

Thurgood Marshall North East

Organizer(s): *Cristiana Mayer, J&J*
Instructor(s): *Frank Bretz, Novartis; José C. Pinheiro, J&J*

The revolutionary advances in basic biomedical science that occurred over the past decade have, so far, failed to translate into comparable improvements in clinical therapies and drugs. In fact, the number of new drug applications (and approvals) has shown a decline over the same period, leading to the so-called "pipeline problem" of the pharmaceutical industry. In response, different initiatives, such as FDA's Critical Path, have been put in place to identify key drivers of poor performance in translating basic science into successful therapies, and to propose ways to address them in clinical drug development. A well-known problem is poor dose selection for confirmatory trials resulting from inappropriate knowledge of dose response relationship (efficacy and safety) at the end of the learning phase of drug development.

This course will discuss, and propose methods to address, the key statistical issues leading to the problems currently observed in dose finding studies, including a review of basic multiple comparisons and modeling methods, as traditionally used in these studies. A unified strategy for designing and analyzing dose-finding studies, denoted MCP-Mod, combining multiple comparison and modeling, will be the focus of the course. MCP-Mod received a positive CHMP qualification opinion in January 2014, as an efficient statistical methodology for model-based design and analysis of phase II dose-finding studies under model uncertainty. It will be discussed in detail, including a step-by-step description of its practical implementation. Case studies based on real clinical trials, together with concrete examples of code in R, will be used to illustrate the use of the methodology. The extension of this framework will be described for count data and time-to-event endpoints and situations involving generalized nonlinear models, linear and nonlinear mixed effects models, and Cox proportional hazards models. A short reference will be made to another extension of the comprehensive multiple comparisons and modeling framework to confirmatory testing in dose-response studies using MCP-Mod.

Short Course 6: Statistical Strategies for Clinical Development of Personalized Medicines

Thurgood Marshall South West

Organizer(s): *Cong Chen, Merck*
Instructor(s): *Cong Chen, Merck*

The future of oncology drug development lies in identifying subsets of patients who will benefit from particular therapies, using putative predictive biomarkers. These technologies offer hope of enhancing the value of cancer medicines, and reducing size, cost, and failure rates of clinical trials. However, inappropriate use of the biomarkers adds cost, complexity, and time to drug development. This short course presents advanced statistical methodologies and strategies for improving the efficiency of and mitigating the risk in late stage development of personalized medicines.

The first part of the short course is devoted to the conventional development paradigm. We will present methods to optimize the design of Phase 2 studies, and adaptively integrate predictive biomarkers into Phase 2 and Phase 3 clinical programs in a data driven manner in which these biomarkers are emphasized in exact proportion to the evidence supporting their clinical predictive value. The resulting program is designed to optimally harvest the value from predictive biomarkers. The second part of the short course is devoted to the expedited development paradigm in that Phase 3 randomized confirmatory trials are initiated at risk after significant preliminary anti-tumor activities are observed in small Phase 1/2 single arm studies. We will present an informational design strategy for risk mitigation. The strategy is applied to address a wide range of issues including de-selection of non-performing biomarker subpopulations.

Students taking this short course will have an opportunity to learn the relevant state-of-art statistical techniques, exchange ideas and immediately apply the learning to practice.

Short Course 7: Bayesian Adaptive Phase I Oncology Trials: Methodology and Implementation

Madison AB

Organizer(s): *Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals*
Instructor(s): *Beat Neuenschwander, Novartis; Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals*

Phase I trials in oncology are usually small adaptive dose-escalation trials. The aim is to approximately understand the dose-toxicity profile of a drug, and, eventually, to find a reasonably safe dose for future testing. A lot of statistical research for Phase I trials has accumulated over the past 25 years, with modest impact on statistical practice. The vast majority of trials still follow the 3+3 design, despite the fact that it often misses the targeted dose (poor operating characteristics) and fails to provide a real understanding about true toxicity rates (no statistical inference).In this course we present a comprehensive and principled statistical approach. The implementation is Bayesian, with the following main parts: a parsimonious model for the dose-toxicity relationship; the possibility to incorporate contextual information ("historical data") via priors; and, safety-centric metrics (overdose probabilities) which inform dose adaptations under appropriate overdose control.

After some basic clinical and statistical considerations, we introduce the statistical methodology for the single-agent setting, and then extend it to dual- and triple-combinations. Applications and a discussion about implementation (such as basic WinBUGS code) issues complement this training and provide practical insights into Phase I trials.

Short Course 8: Designing Observational Comparative Studies Using Propensity Score Methodology in Regulatory Settings

Wilson ABC

Organizer(s): *Nelson Lu, FDA/CDRH; Yunling Xu, FDA/CDRH*
Instructor(s): *Donald Rubin, Harvard University; Lilly Q Yue, FDA/CDRH*

Although well-controlled and conducted randomized clinical trials (RCT) are viewed as gold standard in the safety and effectiveness evaluation of medical products, including drugs, biological products and medical devices, observational (non-randomized) comparative studies play an important role in medical product evaluation, due to ethical or practical reasons, in both pre-market and post-market regulatory settings. However, various biases could be introduced at every stage and into every aspect of the observational study, and consequently the interpretation of the resulting statistical inference could be of concern. Among existing statistical techniques for addressing some of the challenging issues, propensity score methodology is one increasingly used in regulatory settings, due to its unique future of separating "study design" and "outcome analysis."

This course will introduce the causal inference framework and propensity score methods (e.g., matching, stratification, and weighting), and highlight the principle and importance of prospective design of observational comparative studies to increase the integrity and the interpretability of outcome analysis results. Practical issues encountered in the application of the methodology in the regulatory settings will be presented, including but not limited to study design process in regulatory submissions, specification of treatment effects of interest in treatment comparisons (average treatment effect (ATE) or average treatment effect on the treated (ATT)), covariate identification and inclusion, control group selection/

formation (a concurrent control, historical control or a control group extracted from national/international registry), sample size and power consideration. Some differences for implementing propensity score methodology will be delineated for studies with different purposes, for regulatory submissions or general comparative effectiveness research. For example, exclusion of treated patients with an investigational product should be discouraged in studies aimed at pre-market regulatory submissions. These topics will be illustrated with examples based on regulatory review experience.

PLENARY SESSIONS

Thursday, September 17
8:15 a.m. – 9:45 a.m.

The Future of Precision Medicine
Organizer(s): *Richard Zink, JMP Life Sciences, SAS Institute; Wei Zhang, FDA/CVM*

Precision Medicine Initiatives at FDA
Lisa LaVange, FDA

In this collaborative presentation, I will describe various initiatives ongoing at FDA in personalized or precision medicine. Examples based on our experience with targeted therapies, enrichment trial designs, and master protocols will be provided. The statistical reviewer's involvement in these initiatives is important and will be discussed. Future directions for personalized medicine and their applicability in a regulatory setting also will be discussed.

Micro-Randomized Trials & mHealth
Susan Murphy, University of Michigan

Micro-randomized trials are trials in which individuals are randomized 100s or 1,000s of times over the course of the study. The goal of these rials is to assess the impact of momentary interventions (e.g., interventions that are intended to impact behavior over small time intervals). A fast-growing area of mHealth concerns the use of mobile devices for both collecting real-time data, for processing this data and for providing momentary interventions. We discuss the design and analysis of micro-randomized trials for use in mHealth.

Thursday, September 17
10:00 a.m. – 11:30 a.m.

Panel Discussion on the Future of Precision Medicine
Organizer(s): *Richard Zink, JMP Life Sciences, SAS Institute; Wei Zhang, FDA/CVM*
Panelist(s): *Greg Campbell, formerly of FDA/CDRH; Cong Chen, Merck; Lisa LaVange, FDA; Susan Murphy, University of Michigan; Estelle Russek-Cohen, FDA/CBER; Richard Simon, National Cancer Institute*

PARALLEL SESSIONS

Thursday, September 17
1:15 p.m. – 2:30 p.m.

PS1A

Statistical Experiences on Subgroup-Stratified, Biomarker-Stratified, or Enrichment Trials
Thurgood Marshall North

Organizer(s): *Eva Miller, inVentiv Health Clinical; Liming Dong, FDA; Yuanjia Wang, Columbia University; Xiting (Cindy) Yang, FDA/CDRH*
Chair(s): *Xiting (Cindy) Yang, FDA/CDRH*

In a situation that a biomarker/baseline characteristic is identified with great potential in predicting treatment effect of a new therapy, a stratified design may be used: All patients are randomly assigned regardless of subgroup status, but the analysis plan is centered on testing treatment effect dependence on subgroup status. Similar situation applies to trials with concerns on heterogeneity in treatment effect among subgroups defined by important baseline characteristics such as age and gender. At other times, a biomarker is not known at the beginning of a study and an analysis plan may include an identification of the biomarker, followed by an adaptive enrichment plan. In this session, speakers from academia, industry and government will share their experience on these.

Opportunities of Enrichment Designs in the Era of Precision Medicine
Bo Huang, Pfizer Inc.

Traditional clinical development of an experimental therapy utilizes the "one-size-fits-all" approach by testing the regimen in an unselected or untargeted patient population with a specific disease. The assumption is that response in the population with the disease is homogeneous. With the advent of targeted therapies, selection of treatment can be tailored to the genetic makeup of each individual. Therefore, these targeted therapies may benefit only a subset of the entire population and traditional statistical designs may no longer be appropriate or efficient. Statistical designs involving predictive biomarkers generally fall into 2 categories: classical designs and adaptive designs. We give a brief overview of the literature, and discuss the challenges and opportunities in the era of biomarker-based personalized medicine from a pharmaceutical industry perspective with some recent examples and case studies. SAS codes on implementation and evaluation of operating characteristics of some methods will be available in an upcoming book chapter.

Optimal Subgroup Sample Size Allocations in Clinical Studies
Guoxing (Greg) Soon, FDA/CDER/OTS/OB

Subgroup analysis is a common practice in clinical research and drug evaluation. Subgroup analysis by gender, race and age are routinely required by FDA for each NDA submission. The purpose of subgroup analysis is primarily for assessing internal consistency of treatment effect, signaling high-risk subpopulation, and generating hypothesis for further studies. The importance and limitation of subgroup analysis was discussed in several guidances.

The primary concerns for subgroup analysis are 1) the lack of sufficient pre-planning and multiplicity adjustment, which cause difficulty in result interpretation, and 2) small sample sizes in subgroups limit the ability in making inference in the size and consistency of treatment effects among the subgroups. In this talk, we will address issues related to the second concern by exploring the optimal sample size allocation among subgroups, taking into considerations of the prevalence, prior information, cost and time, as well as potential post marketing confirmation.

Determining the Intended Use Population in Phase III Clinical Trials
Richard Simon, National Cancer Institute

Standard clinical trial methodology provides excellent control of type I error, but is much less adequate for specification of the intended use population. Phase III clinical trials often have very broad eligibility criteria but accrued patients rarely represent a random or representative sample of the eligible population. Consequently use of the eligibility criteria as a basis for specification of the intended use population has little statistical basis. Many classically defined diseases have been found to be heterogeneous mixtures of molecularly distinct entities sharing symptomatology but not pathogenesis or responsiveness to treatment. Consequently, standard clinical trial practice often leads to small average treatment effects, large NNT values and substantial over-treatment of the patient population.

Adaptive enrichment designs enable adjustment of the eligibility criteria in a group sequential manner to focus the trial on the kinds of patients who are demonstrating benefit from the test treatment, thereby increasing statistical power and a improving the specification of the intended use population. This presentation will discuss the specification of the intended use population in adaptive enrichment designs.

PS1B

Current and Future Role of the Clinical Statistician in the World of Data Transparency
Thurgood Marshall South

Organizer(s): *Jeffrey Joseph, Theorem Clinical Research; Stephen Wilson, FDA/CDER/OTS/OB/DBIII; Vivian Shih, AstraZeneca; Yueqin Zhao, FDA*
Chair(s): *Stephen Wilson, FDA/CDER/OTS/OB/DBIII*

Pharmaceutical and biotechnology companies submit large amounts of clinical data and analysis to support approval of their drugs and devices with the expectation for this information to be kept confidential as has been the practice by regulators around the years for decades. With the current pressure mounting, regulators and industry have to review policies and procedures on how they will release this data to enhance the public health. The future will include data transparency. The FDA issued the deliberations from the Transparency Task Force (2011) and the EMA (2013) has released a draft policy document titled *Publication and Access to Clinical Trial Data*.

PhRMA (2011) has indicated that pharmaceutical and biotechnology companies are committed to enhancing the biomedical research through the responsible sharing of clinical trial data, which: protects patient privacy for research participants, preserves the integrity of regulatory systems, and maintains incentives for investments in biomedical research. Such responsible data sharing would enhance the public health and accelerate development of new drugs and devices by allowing reanalysis of the existing data compiled by sponsor supported clinical trials.

The procedures and processes for data transparency to occur with reanalysis of the clinical data to undergo scientific rigor has possible new roles and responsibilities for the statistician. These roles and responsibilities would fall in three categories: 1) access to the clinical data through independent panel, 2) planning and analysis of the data, 3) future planning and design of clinical trials.

For a researcher to gain access to clinical data under data transparency, a number of pharmaceutical companies have setup an independent panel (i.e., GSK has an independent panel (2014), Janssen has Yale Open Data Access (YODA) at Yale University). The roles and responsibilities of the statistician as a member of this panel and the criteria for review of the research proposal plan (design, methods, and analysis) to meet scientific objectives. During this session, a statistician currently in this role will update those attending the session on his/her experience.

A researcher in conjunction with a statistician will need to provide the formal statistical analysis plan to gain access of the data through a gatekeeper, and formal study report will need to be provided within 1 year of completing reanalysis for public access. In this session we would discuss the analysis necessary for multiple clinical studies such as meta-analysis, and any new or exploratory analysis; also, we would have to discuss the issues of multiplicity and sub-group analyses. The topic of generation and combining of patient level data for ease of analysis would also be discussed and would include appropriate methods of de-identification, standardization of the data (e.g., CDISC), and appropriate coding (i.e., medication procedures, adverse events).

Last, this session would address how data transparency would affect the planning and design of future clinical trials. Will there be an effect of the sample size of the future trial? Would this increase or decrease the use of adaptive design in future trials?

Supporting Open Access for Researchers
Michael Pencina, Duke Clinical Research Institute

There are growing calls for increased access and transparency of clinical data. The recent report from the Institute of Medicine (IOM) outlined several key features of this process. In response, the Duke Clinical Research Institute (DCRI) through its new SOAR (Supporting Open Access for Researchers) initiative aims to facilitate open sharing of trial data with interested researchers. SOAR specifically focuses on several key areas underscored in the IOM report, including an independent review committee that ensures expert consideration of each proposal; stringent data de-identification and protection of patient privacy; promotion of pre-specified statistical analysis plans (SAP); and independent review of final manuscript prior to submission. In this presentation we will present details of the SOAR model and comment on several outstanding aspects of the IOM report.

Data Sharing, Year 2: Access to Data from Industry-Sponsored Clinical Trials
Marc Buyse, IDDI Inc.

Since May 2013, investigators have been able to request access to de-identified individual patient data from clinical trials sponsored by GlaxoSmithKline, subject to review and oversight by an independent review panel. The first year of experience with this data sharing initiative was reviewed recently. A wide variety of research projects were granted access to patient level data. The 36 projects approved could be broadly categorized as follows:

• Studies aimed at comparing treatment regimens (n=3; e.g., to perform a meta-analysis of studies of epilepsy treatments)

• Studies aimed at optimizing treatments (n=3; e.g., to calibrate activated clotting time to avoid bleeding complications)

• Studies for patient stratification (n=3; e.g., to assess differential effects across ethnic groups)

• Studies of risk factors or biomarkers (n=6; e.g., to assess the effect of tumor size on survival)

• Methodologic studies (n=5; e.g., to develop predictive toxicology tools)

• Miscellaneous other studies (n=1 each; e.g., to study the relationship between influenza vaccination and lifestyle factors; to estimate the effectiveness of paroxetine in treating depression in adolescents; to elucidate the expected frequency of severe adverse events before designing a new trial to be conducted in the developing world, etc.)

In this talk, I will provide updated information on the data sharing initiative, which has now been joined by several pharmaceutical companies. I will also discuss how this initiative could be extended and improved for maximum benefit to future research and patients.

Yale Open Data Access (YODA): J&J Collaboration with Yale
Jeffrey Gardner, Janssen R&D

The responsible sharing of clinical trial data with researchers external to the sponsor company requires a balance between maintaining patient privacy and researcher access to data. Janssen has partnered with the Yale Open Data Access project to help balance these two elements. There are three basic components to responsible data sharing.

• Request Intake. The first is the collection of the researcher's request, which includes identification of clinical trials and proposed analysis plans. The YODA project website allows researchers globally to request access to Janssen clinical trial data.

• Request Evaluation. Each request received is reviewed for completeness and evaluated for merit with respect to advancing public health knowledge. The independent evaluation is performed by the YODA project, which has access to a steering committee for guidance. Janssen is blinded to the identity of the requester until the request is approved/declined and posted on the YODA project website. Requests to support commercialization or litigation are outside of the scope of the YODA project.

• Request Fulfillment. The sharing of patient level data (PLD) is provided in the form of a safe harbor certified analysis platform. Researchers are granted access to requested trial data and have a variety of analysis tools

available to perform their research. PLD and analysis output remain within the platform until the researcher is ready to publish their results.

PS1C

Big Data/Big Analytics: Challenges and Opportunities in Pre- and Post-Market Medical Product Evaluations Utilizing National/International Registries

Thurgood Marshall East

Organizer(s): *Yunling Xu, FDA/CDRH; Nelson Lu, FDA/CDRH; Charles Darby, Statistical Consultant; Rakhi Kilaru*

Regulatory decisions are made based on the assessment of risks and benefits of the medical products, including drugs, biologics and medical devices, at the time of pre-market approval, and subsequently, when post-market risk-benefit balance needs reevaluation. Such assessments depend on scientific evidence obtained from pre-market studies, post-approval studies, post-market surveillance studies, and relevant registries. Currently, national/international registries are playing more and more important roles in the safety and effectiveness evaluations of medical products, in both pre- and post-market settings. Although such registries provide a huge amount of data reflecting real world practice, challenges arise concerning how to use the data to draw reliable statistical inferences. This session will focus on why and how to prospectively design clinical studies utilizing registries for objective causal inference. Statistical and regulatory challenges and opportunities will be presented with examples. Various issues will be discussed by a panel of statistical and medical experts from academia, industry and government, from both pre- and post-market perspectives.

Objective Observational Study Design Using Big Data for Causal Inference
Donald Rubin, Harvard University

The principles of the objective design of observational studies do not depend on the size of the data sets being studied, but rather follow well-established principles that have evolved from the design of proper randomized experiments. Of course situations involving observational studies are typically more complex than those involving randomized experiments, and larger data sets create more opportunities for innovative and clever analyses as well as opportunities for inapposite data-dredging exercises. The focus of this presentation will be on practical advice to implement the former while avoiding the pitfalls of the latter.

Designing Observational Comparative Studies Using Registry Data in Regulatory Settings
Lilly Yue, FDA/CDRH; Nelson Lu, FDA/CDRH; Yunling Xu, FDA/CDRH

National and international registries, which reflect real world evidence, can play important roles in regulatory decision making in both pre-market and post-market settings. This presentation focuses on developing innovative strategies to utilize such data—prospectively designing clinical studies with propensity score methodology. Examples will be given based on regulatory reviews.

PS1D

Analytical Similarity: Current Statistical Issues in Biosimilar Product Development

Thurgood Marshall West

Organizer(s): *Meiyu Shen, FDA; Harry Yang, MedImmune LLC*
Chair(s): *Harry Yang, MedImmune LLC*

The biosimilars regulatory program has been taking shape at the FDA over the past three years. Several guidance documents have been issued. In parallel, associated statistical methodologies have been evolving. For example, statistical accommodation of limited availability of biological materials and lots has been developed. Statistical assessment of biosimilars in a clinical setting also remains an outstanding issue as we grapple with the new challenges arising from different endpoints and limited information. This session consists of two presentations, one expert statistician representing an industry perspective and one representing a regulatory perspective, who will together provide fresh insight on the application and issues surrounding statistical approaches to analytical similarity. There will be ample time for audience discussion and participation during the second half of the session.

PS1E

DMCs and Adaptive Clinical Trials: Considerations in Balancing Safety and Trial Integrity

Lincoln 5

Organizer(s): *Michelle Detry, Berry Consultants; Jie (Jack) Zhou, FDA/CDRH; Greg Ball, AbbVie; Zhuang Miao, FDA*
Chair(s): *Michelle Detry, Berry Consultants*

Traditionally, data monitoring committees (DMCs)—the independent bodies tasked with overseeing ongoing clinical trials with the goals of mitigating risks to subjects and protecting the scientific integrity of the trial—have been given substantial latitude in the scope of the information they consider and the types

of recommendations they may make. FDA's thoughts regarding the function of DMCs were captured in a March 2006 FDA guidance document titled *Guidance for Clinical Trial Sponsors: Establishment and Operation of Clinical Trial Data Monitoring Committees* (see *www.fda.gov/downloads/Regulatoryinformation/Guidances/ucm127073.pdf*).

Since the issuing of the FDA guidance on DMCs, there has been increasing interest in the use of adaptive clinical trials—trials in which key aspects of the trial are modified during the trial in response to data accumulating within the trial itself, according to pre-specified rules, to achieve goals of efficiency, improved patient outcomes, or better ethical balance—in both the exploratory and confirmatory phases of clinical development. The FDA's current thinking regarding the use of such designs is largely captured in a 2010 draft guidance titled *Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics* (see *www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf*) and the 2015 draft guidance *Adaptive Designs for Medical Device Clinical Studies* (see *www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm446729.pdf*).

When overseeing the conduct of a confirmatory, adaptive clinical trial, a DMC must balance the need for flexibility in responding to unexpected patterns in efficacy and safety and the need to maintain the pre-specified nature of the trial design and statistical integrity for regulatory review. This requires a sophisticated understanding of these competing considerations, regulatory science, and clinical care by the DMC members. As both traditional and innovative adaptive designs have become more prevalent, previously unconsidered issues have arisen, requiring DMC's roles and responsibilities to evolve. This session will use the existing FDA guidance documents and speaker expertise with both participation on, and support of DMCs to illustrate and explore the potential issues and propose solutions.

The objectives of this session are: (1) to briefly review the current FDA guidance document on the responsibilities and operation of DMCs, with particular emphasis on those areas that potentially impact the oversight and integrity of confirmatory or pivotal, adaptive clinical trials; (2) to briefly review the current FDA draft guidance documents on adaptive design clinical trials, with particular emphasis on those areas more likely to impact the roles, responsibilities, and operation of DMCs; (3) to identify and discuss areas of agreement, potential conflicts, and gaps in these FDA guidance documents for guiding the oversight of confirmatory adaptive design clinical trials; (4) to suggest operational procedures and clarifications of roles and responsibilities to be included in DMC charters that best address issues identified above; and (5) to identify gaps in regulatory science related to the oversight of confirmatory, adaptive design clinical trials.

Adaptive Clinical Trials and Data-Monitoring Committees: One View from the World of Medical Devices
Greg Campbell, Consultant

The FDA guidance document *Data Monitoring Committees* was finalized in 2006. In the meantime, the importance, use and experience associated with adaptive designs have grown enormously. FDA has issued two draft guidance documents on adaptive designs, one for drugs and biologics in 2010 and another, more recently, for medical devices in 2015. While the 2006 guidance on DMCs carefully lays out the principles for the establishment and operation of data monitoring committees for clinical trials, this talk will briefly review the 2006 document on DMCs in terms of the potential impact on adaptively designed clinical trial trials. The draft guidance on adaptive designs for medical devices will be reviewed in terms of those areas likely to impact the roles, responsibility and operation of DMCs. Possible changes to the DMC document to reflect the challenges that adaptive designs can make to the operation and responsibility of DMCs will be considered.

DMCs and Adaptive Clinical Trials: Considerations in Balancing Safety and Trial Integrity
Thomas Cook, University of Wisconsin School of Medicine and Public Health

DSMBs play an important role in protecting the safety of trial participants while preserving the scientific integrity of the trial. This requires that they be able to balance potential safety concerns with the potential for the trial to demonstrate benefit. Thus, DSMBs must receive summaries of evolving results that are maximally informative and easy digest. Trials using sophisticated computational algorithms for making adaptations, incorporate an additional layer of complexity. To the extent that the DSMB is directly involved with the adaptations, these should be in addition, rather than in place of their normal data monitoring activities. Furthermore, DSMB decisions must involve the totality of evidence and require the flexibility to make recommendations independent of the adaptive procedures. This requires that the adaptive procedures accommodate for some degree of DSMB discretion without jeopardizing the operating characteristics of the trial procedures. This talk will discuss some the competing interests that DSMBs face and make recommendations regarding how these interests can be accommodated.

PS1F

Advanced Multiple Testing Methodologies for Confirmatory Trials

Lincoln 6

Organizer(s): *Xuan Liu, AbbVie; Freda Cooner, FDA/CDER; Anthony Rodgers, Merck; Julia Jingyu Luan, FDA/CDER*
Chair(s): *Xuan Liu, AbbVie*

There has been considerable development in advanced multiple testing methodologies in recent years in respond to the demand of more efficient confirmatory trial designs. An ideal trial design needs to be tailored not only to meet the sponsor's objectives with "optimal" efficiency but also to ensure valid interpretation of the study results for regulatory considerations. Very often nowadays, such trial design may inherit a very complex multiple testing problem with multiplicities from different sources. For example, a confirmatory trial can include multiple endpoints, multiple doses of the investigational drug, multiple control arms, or multiple populations. It may also include adaptive features such as group sequential designs with early stopping for efficacy or futility, seamless phase II/III designs with adaptations in-between phases, or enrichment designs to narrow down the patient population. The multiple testing strategy is not only one of the most important components in the trial design from the sponsor's perspective, but also an important factor in regulatory agencies' review processes. In this session, a group of prominent researchers from pharmaceutical industry and regulatory agencies will present their research and views on this important topic.

Innovative Clinical Trial Designs That Control for Multiplicity

Walter Offen, AbbVie

In recent years, there have been many discussions at scientific meetings and in the literature regarding innovative clinical trial designs, including adaptive designs and gate-keeping strategies. Many of these present challenges in controlling the Type I error rate. This talk will provide an overview of important designs and related approaches for multiplicity adjustment that improve statistical power. In the event that the FDA Multiplicity Draft Guidance is released prior to this meeting, we will also highlight the content that has the greatest impact on industry-sponsored clinical trials, and will provide commentary regarding the corresponding approaches proposed in that guidance.

Generalized Error Rates for Subgroup Analyses

Frank Bretz, Novartis; Willi Maurer, Novartis; Xiaolei Xun, Novartis

We consider the problem of comparing the treatment effect of a new drug against a comparator for two non-overlapping subgroups of patients defined by predictive biomarkers, demographic factors or any other classifier. A decision is to be made if and for which of the two subgroups the respective null hypotheses can be rejected and an advantage of the new drug over the comparator be claimed. We argue that in this situation traditional methods to control the Type I error rate are too restrictive and that the standard familywise error rate (FWER) is not appropriate. Instead, we propose decision procedures that allow us to control the FWER, but for which also upper bounds of expected values for more general loss functions can be derived.

Validity of the Hochberg Procedure Revisited for Clinical Trial Applications

Mohammad Huque, FDA/CDER

There is much interest in using the Hochberg procedure (HP) for tests on primary endpoints of confirmatory clinical trials. The procedure is simple-to-use and enjoys more power than the Bonferroni and the Holm procedures. However, the HP is not assumption free like the other two procedures. It controls the familywise Type I error rate (FWER) when test statistics (used for statistical tests) are independent or if dependent satisfy a conditionally independent formulation. Otherwise, its properties for dependent tests at present are not fully understood. Confirmatory trials for statistical tests normally use simple test statistics, such as the normal Z, student's t, and chi-square. The literature does include some work on the HP for dependent cases covering these test statistics, but concerns remain regarding its use for confirmatory trials for which endpoint tests are mostly of the dependent kind. The purpose of this presentation is therefore to revisit this procedure and provide some clarity for better understanding of its performance for dependent cases.

Thursday, September 17
2:45 p.m. – 4:00 p.m.

PS2A

Logical Inference on Treatment Efficacy in Subgroups and Their Combinations in Personalized Medicine Development

Thurgood Marshall North

Organizer(s): *Ying Ding, University of Pittsburgh; Xiang Ling, FDA/CDER; Jason Hsu, Eli Lilly and Company/The Ohio State University; Thomas Birkner, FDA*
Chair(s): *Jason Hsu, Eli Lilly and Company/The Ohio State University*

In personalized medicine development, the patient population is thought of as a mixture of two or more subgroups that may derive differential treatment

efficacy. In order to find the right patient population for the treatment to target, it is necessary to infer treatment efficacy in subgroups and combinations of subgroups. A fundamental consideration in this inference process is that the logical relationships between treatment efficacy in subgroups and their combinations should be respected (for otherwise the consistency assessment of efficacy may become paradoxical). Surprisingly, this basic principle is violated by several commonly used efficacy measures and/or popular inference procedures, causing illogical conclusions. In this session, new methods of inference procedures that preserve the logical relationships on appropriately defined efficacy measures will be presented. Presentation by Yi Liu will contrast the different objectives of developing new drugs that target subgroups of patients versus individualizing selection of treatment among existing drugs for each patient. Then, for developing new drugs, she will illustrate how to logically infer on the efficacy of a treatment for patients with biomarker values above a threshold, and how to choose a threshold for the biomarker. Presentation by Ying Ding will show, for time-to-event outcomes and ordinal biomarkers, how to analyze subgroups and their mixtures for treatment efficacy with suitable efficacy measures. James Hung will be a discussant of the presentations.

Thresholding of a Companion Diagnostic Test Confident of Efficacy in Targeted Population
Yi Liu, Takeda Pharmaceuticals

There are two different perspectives on personalized medicine. One perspective is to optimize selecting a treatment for each individual patient among existing treatments, on average. The other perspective, the setting of this presentation, is to develop a new drug that would provide new benefit to a subgroup of patients. That is, patients in this subgroup will do better under this new treatment on average than under the Standard of Care. We will show how to correctly infer efficacy in the subgroup of patients whose biomarker value is above a certain threshold while maintaining the logical relationship among the parameters. We will also show how to find the "optimal" threshold from a new drug development perspective.

Logical Inference on Treatment Efficacy in Subgroups and Their Mixture, with an Application to Time-to-Event Outcomes
Ying Ding, Department of Biostatistics, University of Pittsburgh

In the new drug development process, how to correctly assess treatment efficacy in subgroups and their combinations can be nontrivial. It depends on the nature of efficacy measure as well as the estimation procedure. The current statistical practice of estimating the treatment efficacy in a mixture population has serious flaws. We propose a subgroup mixable estimation principle that respects the logical relationships between treatment efficacy in subgroups and their combinations. Focusing on the time-to-event outcomes and ordinal biomarkers, we develop a simultaneous inference procedure, with appropriate efficacy measures, to correctly infer treatment efficacy in a mixture population.

PS2B

Subgroup Analysis Under Rising Regulatory Emphasis: Fundamentals and Challenges
Thurgood Marshall South

Organizer(s): *Yijie Zhou, AbbVie; Bo Yang, AbbVie; Weiya Zhang, FDA; Yifan Wang, FDA*

Investigation and interpretation of the findings of subgroup analysis in confirmatory clinical trials has always been important and yet challenging. Recently in 2014, initiatives have been undertaken by major regulatory agencies regarding subgroup analysis: EMA issued a draft guideline, collected industry feedback and held a workshop for further discussion; FDA held a public hearing on demographic subgroups and afterwards issued an action plan. With the rising regulatory emphasis on this topic, we will re-convey what are the fundamental statistical components embedded in subgroup analyses that will enable correct decision making regarding subgroups, and how we can address these components in the regulatory environment nowadays.

Issues Related to Subgroup Analysis and Possible Ways for Improvements
Lu Cui, AbbVie; Shufang Liu, AbbVie

Confirmatory randomized clinical trials are designed to provide definitive information on the efficacy and safety of a new drug. While the outcome based on the overall population is the basis for the final conclusion, sound subgroup outcomes may provide additional insights of the drug effects and information for possible development of personalized treatment strategies. This presentation is to revisit issues related to subgroup analysis, to highlight pitfalls of potential biased interpretations and to explore ways for improvements. For the latter, the use of extensive stratifications to reduce the chance of biased subgroup findings is discussed with simulation results to illustrate the idea.

Subgroup Analysis
Janet Wittes, Statistics Collaborative

Subgroups have long been the bane of many statisticians in clinical trials. We cringe when we hear clinicians say, "I treat patients, not means." We, like they, know that different patients respond differently to the

same drug, but we do not know how to predict reliably who will, or won't, respond. So we demand strong evidence of differential subgroup effects to conclude that a subgroup is more (or less) responsive than other subgroups. We warn about the landscape of clinical trials that chased, unproductively, a promising subgroup. But in the light of global trials and targeted therapies, is it time to reassess our skepticism? Or, more narrowly, are there situations where our traditional stance is not useful? This talk provides an overview of methods that aim to identify more reliably subgroups with responses to therapy that differ materially from each other.

PS2C

Large Trials for Major Adverse Cardiovascular Events

Thurgood Marshall East

Organizer(s): *Aloka Chakravarty, FDA/CDER; Bret Musser, Merck; Olga Marchenko, Quintiles*
Chair(s): *Richard Zink, JMP Life Sciences, SAS Institute*
Panelists: *Aloka Chakravarty, FDA/CDER; Qi Jiang, Amgen; José Pinheiro, J&J; Estelle Russek-Cohen, FDA/CBER*

This session will discuss the design and operational aspects of trials with safety objectives, such as the cardiovascular outcome trials (CVOTs) of type 2 diabetes mellitus (T2DM) programs. CVOTs provide an opportunity to assess rare safety signals and better evaluate benefit-risk profiles, but present challenges in statistical and operational areas including efficiency of statistical design, study interpretation, long-term patient retention, data confidentiality and high cost. In this session, experts from the pharmaceutical industry and the FDA will share their thoughts on CV risk assessment strategies in T2DM development programs, and discuss lessons learned and best practices.

The session will open with a presentation given by Olga Marchenko from Quintiles, who is co-chair of the ASA Biopharmaceutical Section Safety Working Group, followed by a panel discussion featuring thought leaders from the biopharmaceutical industry and FDA. During the panel discussion, preplanned questions will be addressed first, followed by questions from the audience. The panel discussion will complement the presentation.

Overview of Strategies for Assessing CV Risks of T2DM Treatments and Arising Questions

Olga Marchenko, Quintiles

This presentation will be based on the work of the ASA Biopharmaceutical Section Working Group on Safety. In 2008, the U.S. FDA released the guidance for industry, *Diabetes Mellitus – Evaluating Cardiovascular Risk in New Anti-Diabetic Therapies to Treat Type 2 Diabetes*, that changed the way new non-insulin anti-diabetes drugs are evaluated and brought to market. With representatives from different institutions, the group reviewed treatments approved by the FDA to treat type 2 diabetes mellitus during 2002-2014 with a focus on cardiovascular (CV) risk assessment. To meet guidance requirements on CV risk assessment, different strategies that include meta-analyses and stand-alone cardiovascular outcome trials (CVOTs) have been conducted. CVOTs provide an opportunity to evaluate safety signals beyond CV risk and better assess the benefit-risk profile in diabetic patients with a high risk for CV events, but they also present numerous challenges. The advantages and disadvantages of different CV assessment strategies will be summarized, and some emerging questions will be raised in the presentation.

PS2D

Quality and Quality Metrics

Thurgood Marshall West

Organizer(s): *Stan Altan, J&J; Liang Zhao, FDA*

The Food and Drug Administration Safety and Innovation Act (FDASIA) of 2012 gave the FDA broad additional authority including developing a uniform set of standards for all regulated products, generic, brand name and OTCs. The concept of quality metrics is in the process of being articulated and promoted by the new office of pharmaceutical quality. A guidance on quality metrics is in progress. FDA has emphasized the importance of a company's quality culture and the intent to identify and measure this culture with corresponding metrics. In May 2014, Russ Wesdyk of FDA proposed four "consensus quality metrics" with definitions and stated that additional metrics were still being defined. In June 2014, ISPE initiated an industry-wide quality metrics pilot through McKinsey to collect data for calculation of the "four consensus metrics" plus others that were defined by the ISPE team. A published summary of ISPE's quality metrics pilot data analysis and learnings is anticipated for further discussion at the ISPE Quality Metrics Summit in April 2015. The FDA-Industry workshop would dedicate one session to this important topic of "quality" metrics being developed by the FDA to promote a dialogue on the impact they might have on statistical practice related to the assessment and improvement of quality at biopharmaceutical companies.

Quality Metrics: Why It Matters

Lawrence Yu, FDA/CDER/OPQ

The FDA has within the past month issued a draft guidance on quality metrics. To provide a deeper understanding of the concepts embodied in the

draft guidance, I will discuss the importance of QbD concepts to the evolution of quality metrics, the FDA rationale for its development and expand on the OPQs vision intended to enhance the role of a quality culture within industry.

1. Product quality in relation to QbD
   a. QbD principles and objectives
   b. QbD during development and commercialization
2. FDA rationale for the development of quality metrics
   a. Overview of draft quality metrics guidance
   b. Relationships between the cGMP and quality metrics
3. Quality evolution

Process Capability: Is It Just a Quality Metric?
Helen Strickland, GSK

In the first half of the 20th century, W.A. Shewhart and W.E. Deming promoted the view that the long-range contribution of Statistics depends not so much upon getting a lot of highly trained statisticians into industry as it does in creating a statistically minded generation of physicists, chemist, engineers and others who will in some way have a hand in developing and directing the production processes of tomorrow. Well into the second decade of the 21st century, the contribution of statistics in the pharmaceutical manufacturing industry is still evolving. This is evidenced by the multitude of regulatory guidance documents that advocate the use of statistical process management tools to support the implementation of quality management systems. quality metrics, such as lot acceptance rate, product quality complaint rate, invalidated out-of-specification rate have been identified by the FDA as being valuable in assessing the overall effectiveness of a company's quality management system. The appropriateness of process capability indices (PCIs) is also being considered as PCIs are a measure of the process's ability to produce product that complies with specifications. Most individuals jump right to the computation of the PCIs, thereby placing more value on the metric itself than the product and process knowledge obtained through the appropriate evaluation of the manufacturing process. This presentation initiates a discussion of how to appropriately use the information obtained from process capability/process performance assessments.

PS2E

Current Statistical Issues in Biosimilar Product Development

Lincoln 5

Organizer(s): *Bo Jin, Pfizer Biotechnology Clinical Development; Sungwoo Choi, FDA/CDER; Jason Liao, Novartis; Xin Gao, FDA/CDER*
Chair(s): *Joshua Chen, Sanofi Pasteur; Eric Chi, Amgen Inc.*

There have been a few FDA draft guidance on the development of biosimilar products. Specifically, the FDA draft guidance in 2012 on scientific considerations in demonstrating biosimilarity to a reference product recommends that sponsors use a stepwise approach in their development of biosimilar products and indicates that FDA considers the totality of the evidence provided by a sponsor to support a demonstration of biosimilarity. General scientific principles are discussed in the guidance on conducting comparative structural and functional analysis, animal testing, human PK and PD studies, clinical immunogenicity assessment, and clinical safety and effectiveness studies. The other FDA draft guidance in 2012 on quality considerations provides some recommendations on the scientific and technical information of the chemistry, manufacturing, and controls (CMC) section of a marketing applicant for a proposed biosimilar product. The FDA 2014 draft guidance on clinical pharmacology data to support a demonstration of biosimilarity to a reference product further discusses some concepts related to clinical pharmacology testing for biosimilar products and the approaches for developing the appropriate clinical pharmacology database, and the utility of modeling and simulation for designing clinical trials. Despite these efforts, there still remains to be a number of statistical questions to be answered for the regulations and the development for biosimilar products. This includes, but is not limited to, standardization of manufacturing quality control, assessment of variability, stability testing and quality comparison, methodology to demonstrate clinical PK and PD similarity, statistical considerations for clinical efficacy and safety comparability trials including determination of equivalence margin and statistical assessment of immunogenicity similarity, and statistical study design and assessment of biosimilar interchangeability etc. This session provides an excellent opportunity to statisticians at FDA, academia and the industry to work together and join force to discuss these challenging issues on the development of biosimilar products. The presentations in this session will span all the stages of biosimilar development, from CMC, clinical PK and PK/PD, to clinical efficacy and safety comparability trials. The session will consist of three presentations and the discussions with industry, academia and FDA representatives.

Equivalence Test for Two Emax Curves in Biosimilar Studies
Bo Jin, Pfizer Biotechnology Clinical Development; Kerry Barker, Pfizer Inc.

The 2014 FDA draft guidance of clinical pharmacology data to support a demonstration of biosimilarity to a reference product considers clinical pharmacology studies as a critical part in the clinical evaluations on similarity between a potential biosimilar and the reference product. While bioequivalence approach can be applied to PK similarity assessment in clinical pharmacology studies as indicated in the draft guidance, there remains to be a few unsettled problems to evaluate the similarity in terms of pharmacodynamics activities, specifically, including the problem to demonstrate the similarity on response profiles between two products. In this presentation, we describe a new procedure to test equivalence of two E-max curves which can be applied to both time-response and dose-response similarity assessments. Simulations results will be presented to compare the new procedure to other traditional procedures in terms of both Type I error and power performance. Discussions will also be provided on endpoint and dose selections from E-max response profile to provide sensitive assessment on similarity.

Statistical Approaches to Demonstrate Analytical Similarity of Quality Attributes
Cassie Dong, FDA

Compared to drugs, which are usually small molecules, biologic products have much more complicated structures and manufacturing process. Thus, unlike generic drugs where the active ingredients are identical to the reference product, biosimilars are similar to the reference product in terms of quality, efficacy and safety. The development of biosimilars consists of analytical, non-clinical and/or clinical studies. As the fundamental part in the development process, analytical similarity assessment consists of a comprehensive comparison of the physicochemical attributes and biological activities between the biosimilar and the reference products. Statistical equivalence testing plays a critical role in providing quantitative assessment for analytical biosimilarity.

In this talk, we will start with an introduction of analytical biosimilarity assessment, followed by the application of statistical equivalence testing in this process. The data set from Zarxio, the 1st approved biosimilar product in the United States, will be presented as an example.

Statistical Issues in Comparative Clinical Studies of Biosimilars
Gregory Levin, FDA

In this talk, I will focus on the design, conduct, and analysis of the clinical study comparing the safety and effectiveness of the reference product and the proposed biosimilar. I will discuss key issues such as the choice and justification of the similarity margin, endpoint selection and the use of surrogates, prevention and treatment of missing data, and extrapolation of findings from one studied indication to all approved indications.

PS2F

Adaptive Enrichment Design: A Way to Achieve the Goal of Personalized Medicine?
Lincoln 6

Organizer(s): *Min (Annie) Lin, FDA/CBER; Zhiwei Zhang, FDA/CDRH; Feng Liu, GSK; Inna Perevozskaya, Pfizer Inc.*
Chair(s): *Meijuan Li, FDA; Feng Liu, GSK*

Growing interest in adaptive enrichment designs, which involve preplanned rules for modifying enrollment criteria based on accrued data, has been recognized in pharmaceutical researches. An adaptive enrichment study is usually designed to decrease the heterogeneity of patients being studied and therefore considered as a way for pursuing personalized medicine. While most statistical literatures on adaptive enrichment designs focus on the approaches with pre-specified subpopulation characteristics identified prior to or early in clinical development (e.g., a predictive biomarker), some recent methodologies paid more attention to work on adaptively choosing the entry criteria based on interim observations. In this session, we will present the recent developments in methodology and case studies of adaptive enrichment design trials. Discussion will be made to better understand the methodological issues as well as challenges in implementation from statistical, operational and regulatory perspectives.

Adaptive Biomarker Population Selection and Enrichment in Confirmatory Phase III Trials
Cong Chen, Merck; Nicole Li, Merck

Oncology drug developers often decide to initiate Phase III trials at risk after significant preliminary anti-tumor activities are observed in small Phase I/II trials. The preliminary data can hardly provide the much-needed information for selecting a biomarker cutpoint or prioritizing a biomarker hypothesis in Phase III. To address this issue, Magnusson and Turnbull (2013) proposed a group sequential enrichment design that de-selects non-performing biomarker subpopulations at an interim analysis and pools the remaining ones in final analysis. The same endpoint was used for interim and final analyses therein. In this presentation, we propose a more general approach in that different endpoints may be used (e.g., PFS for interim analysis and OS for final analysis) and sample size for remaining ones is subject to increase after the interim analysis. An interesting multiplicity issue will

be discussed. The use of a sensitive intermediate endpoint for population de-selection increases the study power after multiplicity adjustment, which is further improved with sample size adjustment. Our proposed design paves the way for expedited development of personalized medicines in confirmatory trials with limited prior data.

Subgroup Selection in Adaptive Signature Designs of Confirmatory Clinical Trials
Zhiwei Zhang, FDA/CDRH

The increasing awareness of treatment effect heterogeneity has motivated flexible designs of confirmatory clinical trials that prospectively allow investigators to test for treatment efficacy for a subpopulation of patients in addition to the entire population. If a target subpopulation is not well characterized in the design stage, it can be developed at the end of a broad eligibility trial under an adaptive signature design. We propose new procedures for subgroup selection and treatment effect estimation (for the selected subgroup) under an adaptive signature design. We first provide a simple and general characterization of the optimal subgroup that maximizes the power for demonstrating treatment efficacy or the expected gain based on a specified utility function. This characterization motivates a procedure for subgroup selection that involves prediction modeling, augmented inverse probability weighting, and low-dimensional maximization. A cross-validation procedure can be used to remove or reduce any selection bias that may result from subgroup selection, and a bootstrap procedure can be used to make inference about the treatment effect in the selected subgroup. The proposed approach is evaluated in a simulation study and illustrated with a real example concerning human immunodeficiency virus infection. The main ideas of this work generalize easily to other designs that involve data-driven subgroup selection, including adaptive enrichment designs.

Optimal, Two Stage, Adaptive Enrichment Designs for Randomized Trials, Using Sparse Linear Programming
Michael Rosenblum, Johns Hopkins Bloomberg School of Public Health

Adaptive enrichment designs involve preplanned rules for modifying enrollment criteria based on accruing data in a randomized trial. These designs can be useful when it is suspected that treatment effects may differ in certain subpopulations, such as those defined by a biomarker or risk factor at baseline. Two critical components of adaptive enrichment designs are the decision rule for modifying enrollment, and the multiple testing procedure. We provide a general method for simultaneously optimizing both of these components for two stage, adaptive enrichment designs. The optimality criteria are defined in terms of expected sample size and power, under the constraint that the familywise Type I error rate is strongly controlled. It is infeasible to directly solve this optimization problem since it is not convex. The key to our approach is a novel representation of a discretized version of this optimization problem as a sparse linear program. We apply advanced optimization tools to solve this problem to high accuracy, revealing new, optimal designs. This is joint work with Xingyuan (Ethan) Fang and Han Liu at Princeton University.

Thursday, September 17
4:15 p.m. – 5:30 p.m.

PS3A

Concerns with Reanalysis for Ongoing Data Transparency Initiatives

Thurgood Marshall North

Organizer(s): *Theodore Lystig, Medtronic, Inc.*
Chair(s): *Michael Hale, Amgen*

Recent publications (Ebrahim et al 2014, Christakis et al 2013, Krumholz et al 2014) have called attention to an emerging problem of poorly defined practices and conflicting results for reanalysis of previously reported studies. While there are generally accepted practices of pre-specified statistical analysis plans for original studies, there appears to be a notable absence of agreement for standards for reanalyzing existing data. The Christakis et al 2013 viewpoint article discussed some of the situations leading to questionable reanalyses and recommended some practices to address those. Ebrahim (2014) and colleagues reviewed 37 published reanalyses and compared the results and interpretation with those originally reported, finding unacceptably high rates of discordance, even when the reanalyses were conducted by the same people who performed the original analysis. As access to data sets from completed trials from industry, government, and others continues to increase, we have a rapidly increasing potential for confusion for prescribers, payers, and patients, as well as posing difficult labeling and market access questions for regulators. In view of this, reanalysis has important implications for our health care systems, and we should all be concerned that reanalysis delivers on its promise of greater certainty of our understanding of the benefits and risks of therapy.

The speakers will highlight some of the inherent problems of reanalysis by a secondary party, including degradation of source data due to de-identification and masking, working in a glove-box environment for analysis, possible lack of clarity regarding study conduct, measurements, and endpoints, and other complications. Some provocative questions will be posed for further discussion, such as the role of regulators in

evaluating findings from reanalyses, and inappropriate practices related to specification, conduct, and reporting of reanalysis that could lead to bias and misinformation. Early efforts to address these problems and practices will be presented, including potential ways forward.

Reanalysis of De-Identified Data
Jonathan Hartzel, Merck

To protect the privacy of subjects, a necessary step prior to sharing of any clinical data is to de-identify the data. However, the process of de-identification may impact the results found in a reanalysis, or prevent such a reanalysis entirely. Additionally, while there have been some general proposals on the methods to use for de-identification, the actual implementation can (and does) vary across the Pharmaceutical Industry. This may lead to difficulties in combining data across Sponsors and interpreting results. This talk will review the current de-identification processes used in the Pharmaceutical Industry, highlighting their differences and the potential impact on reanalysis and/or new analyses. Suggestions will be made to help minimize the impact of de-identification on such analyses.

Re-Use of Clinical Trial Data: An FDA Perspective
Estelle Russek-Cohen, FDA/CBER

Sponsors of clinical trials submit data to FDA on a regular basis. The data sets are submitted largely in support of a product they wish to market. The data is regarded as the intellectual property of the sponsor but re-use of data is useful in regulatory decision-making.

Meta-analyses can assess consistency across a class of clinical trials. They can also be used to identify endpoints that are possible surrogate endpoints through meta-regression methods. Exploratory analyses are great for science and generate potential hypotheses. A subgroup identified as a result of an exploratory analysis would still need confirmation with additional studies but that doesn't mean researchers shouldn't look. I see tremendous opportunities for advancing our understanding of subgroups, and for precision/personalized medicine if we have access to the actual data used in trials. The main challenge with making data available for re-use by groups outside a regulatory setting is the need to protect patient privacy and to determine if anonymization will impact any of the conclusions. This is up to companies as FDA is still bound to maintain confidentiality but it doesn't mean FDA cannot be supportive of the effort.

Concerns with Reanalysis for Ongoing Data Transparency Initiatives
Sara Hughes, GSK

When a sponsor prepares to conduct the primary analysis on a recently completed study, they have access to the full set of collected data and all relevant supporting documentation. As the statistical programmers operationalize the statistical analysis plan (SAP), they have access to the full study team in the (likely) event some judgment needs to be applied in interpreting the SAP, or in the (even more likely) event data collected during the course of the study requires some special handling. Will an independent analyst with no involvement in the original study design and analysis have insight into some of these judgments that were made?

Many studies in many diseases have multiple analyses over time: one at the primary timepoint, with many other updated analyses after that. Some researchers talk about wanting a "data set of record." What does that mean in this context and what is a reasonable expectation when we look to sponsors to share data that could have multiple cutpoints over many years?

Studies conducted in recent years benefit from greater adherence to standards (and recently, more common standards) as well as much more rigorous data stewardship. This translates to a higher likelihood that shared data will be accurate and complete. Older studies do not have this feature. What are reasonable expectations when it comes to gaining access to older data / study information?

When data are shared with independent researchers, the current best practice requires that the data are de-identified, or anonymized, in order to remove any personally identifiable information (PII) and to obfuscate the data sufficiently well to make it difficult for anyone to re-identify a specific patient using the data that was shared along with other publicly available records. In fact, the process of de-identifying data is currently being performed in a variety of ways. While there is a lot of similarity between the approaches that are being used, there remain differences and these differences could be important when deciding what types of reanalyses are appropriate and how to interpret the results of any reanalyses. Consider a scenario where one sponsor redacts the raw AE terms, but leaves intact all coded terms, including rare events. Now consider another sponsor who wishes to minimize the risk of re-identification and so chooses to redact the raw AE terms, along with any coded terms when they are rare events. A meta-analysis which compares medicines from these two sponsors will likely show one medicine has a higher risk with respect to that rare event; that would almost certainly be misleading. How can a researcher be sure what de-identification

algorithms were used in producing the data sets they were granted access to? Is it acceptable to allow sponsors to have a range of choices for what represents an anonymized data set?

With regard to what level of anonymization is acceptable, what is the role of the environment in which data are shared? If a researcher can only get access to the data in a secure "glove box" environment does this imply that the data can be "less anonymized" than if the researcher is given more open access to data?

The quality of any reanalysis is a function of the quality and completeness of the data and meta-data that inform the reanalysis. In light of this, I would suggest that any independent researcher who "discovers" a new finding in their reanalysis that is inconsistent with the original results should consider initiating discussions with the original sponsor(s) to ensure these new findings are indeed newsworthy and not simply a function of one or more of the possible explanations described in this abstract. Such discussions with the sponsor(s) would not challenge the researcher's independence. Indeed if, after the new findings had been "tested," the researcher continued to believe their findings were newsworthy, it should give the researcher even more confidence to publish their results, allowing for an open scientific discussion. On the other hand, if discussions with the sponsor(s) identified issues with the data or documentation or with the approach that had been taken in the reanalysis, the researcher would then be better informed as they continue their research. And the publication of an errant finding would have been avoided.

This talk will touch on these many points and propose areas where greater collaboration across data sharing bodies would be most useful.

PS3B

Bayesian Subgroup Analysis: Opportunities and Challenges in Unmet Medical Need

Thurgood Marshall South

Organizer(s): *Margaret Gamalo, FDA/CDER; David Ohlssen, Novartis; Helen Zhou; Freda Cooner, FDA/CDER*

Traditionally, clinical trials have primarily been concerned with comparing treatments on an entire population to provide the most reliable data about the effects of treatments. But often it is also important to determine whether there are differential treatment effects on subgroups, if there is potential heterogeneity of treatment effect in relation to pathophysiology, if there are practical questions about when to treat, or if there are doubts about benefit in specific groups which are leading to potentially inappropriate undertreatment.

Many of these challenges and some solutions are discussed in the 2014 European Medicines Agency guideline on investigational subgroups. While this document represents an important step in this complex area, there seems to be a need for better tools that quantify the risks associated with key decisions. Bayesian methods provide a natural framework for balancing the risk of overlooking an important subgroup with the potential to make a decision based on a false discovery.

Due to the need to streamline drug development, in areas of high unmet medical need, recently, subgroup analysis has started to play a different role. For example, the FDA released a draft Guidance on antibacterial therapies for unmet medical need and notes the possibility to use innovative design strategies including Bayesian modeling approaches for assessing subgroup-specific treatment effects trials involving multi-site infections instead of starting with multiple clinical trials in different sites (which requires duplication of regulatory and infrastructure efforts). Basket trials have also gained ground in oncology. This method recruits patients via biomarker status instead of cancer type. After biomarker identification, the patients are divided into multiple study arms (baskets) by cancer type, and the drug's impact is assessed within the separate arms as well as within the study as a whole. In other challenging situations, such as drug development pediatrics, the EMA produced a concept paper on extrapolation of efficacy and safety. The document encourages use of Bayesian methods to extrapolated efficacy from source to target population (e.g., adults to pediatrics).

The appeal of subgroup analyses in these scenarios is undeniable. In this session, talks will showcase Bayesian subgroup analysis (e.g., methodology, trial design considerations, and other innovations) or case examples that include investigations in subgroups.

A Bayesian Approach for Designing Phase 2 Clinical Trials with Rare Tumor Types in Oncology

Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals; Santosh Sutradhar, Novartis

Clinical trial with multiple rare tumor type could be challenging. Complete pooling across all tumor types may lead to inaccurate estimates. On the other hand, stratified analyses ignore potential similarities across different tumor type due to pathway effect and can lead to less efficient estimate. We proposed a Bayesian hierarchical model to borrow information across tumor types in multi-arm Phase 2 oncology trial with rare indications. Due to robust nature the proposed model allows dynamic borrowing of information between groups. This implies more borrowing when the groups are consistent and less borrowing when the groups differ. In this way, the model is a compromise

between the two alternate extremes of either a completely pooled analysis or a separate analysis in each group and provide reasonable strata specific and overall estimate. Design characteristics will be illustrated by data scenarios and simulation.

Clustered Hierarchical Modeling for Identifying Promising Subgroups
Kert Viele, Berry Consultants; Liz Krachey, Berry Consultants

Many medical conditions are heterogeneous. As we evaluate new treatments, we must identify specific subgroups where patients would benefit from the treatment. As the number of groups increase in these studies, we find what has been called "the approaching wall," where pooling across different effects is inappropriate, but analyzing the groups separately is not possible due to limited resources. In these situations, we need clever methods to effectively leverage all available information both across and within subgroups.

Hierarchical models hold promise, as they allow borrowing of information across subgroups (allowing larger effective sample sizes) while recognizing there is variation across subgroups. These models make use of the intuition that the data in a single group may not be conclusive by itself, but repeatedly seeing a common trend across many subgroups may be conclusive when viewed as a whole.

The assumed across group distribution heavily influences the performance of hierarchical models. The commonly used single unimodal distribution often works well, but has difficulties when there are "nugget" groups (a single outlying subgroup) or a small number of distinct clusters of subgroups (for example, where the treatment either works well or does not work at all).

We propose a Dirichlet process based clustering approach to the across subgroup distribution, effectively replacing the single normal distribution with a mixture of normals. This approach more flexibly handles many situations, such as allowing single outlying subgroups to be treated differently from the other groups, or allowing a "half and half" mixture. Many of the advantages of a single normal hierarchical model are retained while providing increased flexibility.

Bayesian Hierarchical Models for Multi-Way Subgroup Problems
Gene Pennello, FDA/CDRH

Estimates of treatment effects within subgroups tend to have too much variation relative to the true treatment effects, hampering clinical interpretation. To remove this unwanted random variation, Bayesian hierarchical models can be used to pull or shrink the within-subgroup estimates toward the overall estimate, with the degree of shrinkage depending on variation between to variation within the subgroups. Most presentations on Bayesian hierarchical models consider the simple one-way model of unstructured subgroups. However, subgroups may be defined by the levels of more than one factor (e.g., age, sex, ethnicity). The factors may be additive or may interact in their modification of treatment effect. In this talk, Bayesian hierarchical models for multi-factor subgroup problems will be presented that exploit factor structure. The Bayesian posterior mean of the difference in treatment effect between subgroups will be shown to be an intuitive linear combination of marginal and interaction contrasts that are shrunk according to evidence for main factor and interaction effects, respectively. An advantage of using such models is that a difference between subgroups defined by two levels of one factor is adjusted for any imbalance between treatment arms in other factors as well as for unwanted random variation due to multiplicity. Time permitting, we also consider Bayesian hierarchical models to extrapolate from adult to pediatric use of a medical product while adjusting for differences in covariate distributions between the two subpopulations. Bayesian hierarchical models rely on the assumption that subgroups are related by exchangeability, that is, any ordering of the subgroup-specific treatment effects is equally plausible. Application of this assumption should be done with care, as it can be conservative or anti-conservative, depending on study objectives.

PS3C

Town Hall Session: Roles of Statisticians in Academia, Regulatory, and Pharmaceuticals Industry

Thurgood Marshall East

Organizer(s): *Yulan Li, Novartis; Guoxing (Greg) Soon, FDA/CDER/OTS/OB; Yanming Yin, FDA/CDER/OTS/OB; Keaven Anderson, Merck Research Laboratories*
Chair(s): *Yulan Li, Novartis; Guoxing (Greg) Soon, FDA/CDER/OTS/OB*
Panelists: *Robert Califf, FDA; Lisa LaVange, FDA; Jeffrey Helterbrand, Roche; Janet Wittes, Statistics Collaborative; Bob Temple, FDA; Tom Fleming, University of Washington; Sonia Davis, The University of North Carolina at Chapel Hill; Ramachandran Suresh; GSK*

In the era of personalized medicine with rapid evolving science and technologies across multiple disciplines, the need to stress "roles of statisticians" has apparently become a pertinent and pressing issue that has not been adequately addressed in the realm of "roles of statistics." Incorporating and anticipating scientific

and technological advances and the ability to bridge statistics to other fields are necessary for successful collaborations of statisticians with medical professionals and scientists.

There is no doubt that further development in analyses and methodologies is both fundamental and critical for practicing statisticians. This single aspect alone is far from sufficient to ensure successful collaboration and leadership in multidisciplinary environment including not only project teams, regulatory agencies, but also medical and patient communities.

Developing the ability to broaden advancing scientific knowledge and to incorporate new technologies is of parallel importance to developing communication and interpersonal skills.

Being able to properly blend and apply these skills in a particular context of research while taking into considerations of various practical, regulatory and ethical issues is a "statistical art" that is at the core of "roles of statisticians."

This "statistical art" can take various forms, which encompass demands of statisticians "asking the right questions," "making investigators confront their own assumptions," "using both statistical reasoning and common sense," "having ability to bridge between statistics and other disciplines," "showing leadership in driving the integration of science and technology advances with innovative clinical development," etc. Playing the roles well also motivate the development of innovative and relevant statistical theory and methodology.

In this town hall meeting, we aim to broaden our own view of the roles of statisticians as well as to learn the viewpoints held by medical research and public health personnel in this new era of personalized medicine. We will investigate barriers that prevent statisticians from performing more effectively, explore the journey of successful statisticians, and present some examples of critical roles of statisticians. The session will include a panel of well-known statistical and medical experts who will speak from their own experiences to address these issues.

Role of Statistician in the Era of Personalized Medicine: Perspectives from Academia, Regulatory, and Pharmaceuticals Industry
Janet Wittes, Statistics Collaborative

Panel discussion on various questions from panelist members Robert Califf, FDA; Lisa LaVange, FDA/CDER; Jeffrey Helterbrand, Roche; Janet Wittes, Statcollab; Tom Fleming, University of Washington; Robert Temple, FDA; Sonia Davis, The University of North Carolina; Ramachandran Suresh, GSK

Future Evolution of the Statistician's Role in Drug Development: What Qualifies as 'Good Statistical Practice'?
Robert Califf, FDA

In this mini-presentation, Robert Califf discusses the evolution of good statistical practice from clinical perspectives.

Role of Statisticians on Both Sides of Drug Development
Lisa LaVange, FDA

In this mini-presentation, Lisa LaVange highlights the complementary roles of statisticians from regulatory agencies and industry.

Core Competencies for Statisticians in the Pharmaceutical Industry
Jeffrey Helterbrand, Roche

In this mini-presentation, Jeffrey Helterbrand describes the skills necessary to become a successful statistician in the industry.

PS3D

Continuing Discussion: Statistical Considerations for Continuous Manufacturing Processes

Thurgood Marshall West

Organizer(s): *David Christopher, Merck; Yi Tsong, FDA/CDER; Helen Strickland, GSK*

This is a continuation of the session from last year, extending discussion around the evolving topic of statistical considerations for continuous manufacturing. In addition to presenting an update on the current FDA perspective on this topic, the session will introduce perspective and experience from a non-pharmaceutical industry in which continuous manufacturing processes have been successfully used for many years.

PS3E

New Statistical Methods for Risk Assessment

Lincoln 5

Organizer(s): *Zhiwei Zhang, FDA/CDRH; Mei-Ling Lee, University of Maryland; Mengdie Yuan, FDA; Greg Ball, AbbVie*
Chair(s): *Zhiwei Zhang, FDA/CDRH*

An important aspect of medical treatment evaluation is quantifying the risks of undesirable events, such as adverse events, disease progression or death. Risk-related questions arise in different contexts

(e.g., pre-market evaluation of safety/efficacy/effectiveness, post-market surveillance), and the relevant information may be collected as binary data, count data, time-to-event data, or a combination of several data types. It is important to keep the scientific context in mind in choosing a risk measure and an appropriate statistical method for a given application. This session gives several examples of newly developed statistical methods for risk assessment that are targeted at specific scientific questions.

An Extension of Likelihood Ratio Test-Based Method for Signal Detection in a Drug Class with Application to FDA's AERS Database
Yueqin Zhao, FDA

A likelihood ratio test (LRT), recently developed for the detection of signals of adverse events (AEs) for a drug of interest in the FDA Adverse Events Reporting System (FAERS) database, is extended to detect signals of AEs simultaneously for all the drugs in a drug class. This extended LRT, based on Poisson model (Ext-LRT) and zero inflated Poisson model (Ext-ZIP-LRT) are discussed. Simulation studies are performed to evaluate the performance characteristics of Ext-LRT and Ext-ZIP-LRT as well as their power and sensitivity. The proposed methods are applied to the Gadolinium drug class in FAERS database.

Rigorous Risk Assessment of Patient Subgroups in Late Phase Drug Development
Lei Shen, Eli Lilly and Company

In the development of a new therapy, it is common that late phase trials offer the first opportunity for in-depth investigation of potential safety issues of the new therapy, given the larger number of patients being studied in these trials than in early phase trials. There is often, if not always, much interest in identifying subgroups of patients with elevated risk, as evidenced by large number of pre-specified as well as post hoc analyses regarding subgroups routinely performed by both trial sponsors and regulatory reviewers. In addition, these analyses frequently form the basis for warnings for, or even restrictions in, patient populations within drug labels. We believe that both prospective and retrospective analyses play useful roles in risk assessment of patient subgroups, and recent developments in statistical methodology can improve both types of analyses. In this presentation, I will describe a framework to systematically utilize both prospective and retrospective risk assessment in late phase trials. Specifically, the proposed approach seeks to leverage new statistical methods to address multiplicity issues and effectively identify patient subgroups of interest. Our goal is to perform rigorous risk assessment that enables proper interpretation of findings and sensible decision-making.

Threshold Regressions Models, with Application in a Multiple Myeloma Clinical Trial
Mei-Ling Lee, University of Maryland

Cox regression methods are well known. It has, however, a strong proportional hazards assumption. In many medical contexts, a disease progresses until a failure event (such as death) is triggered when the health level first reaches a failure threshold. I'll present the Threshold Regression (TR) model for patient's latent health process that requires few assumptions and, hence, is quite general in its potential application. We use TR to analyze data from a randomized clinical trial of treatment for multiple myeloma. A comparison is made with a Cox proportional hazards regression analysis of the same data.

PS3F

Practical Experiences Using Meta-Analysis
Lincoln 6

Organizer(s): *Anna Nevius, FDA/CVM; Steven Radecki, ASA; Ed Luo; Lisa Rodriguez, FDA/CVM*
Chair(s): *Virginia Recta, FDA/CVM*

FDA has incorporated the use of systematic review and meta-analysis methods for evaluation of drug safety and effectiveness. Most of the use of meta-analysis at FDA has been for safety evaluations. The Center for Veterinary Medicine (CVM) has successfully employed meta-analysis for establishing substantial evidence of effectiveness FOLLTROPIN, a follicle-stimulating hormone (FSH).

Systematic Review, Meta-Analysis, and the Regulatory Process
Laura Hungerford, FDA/CVM

Approval of new animal drugs is based on evaluation of scientific evidence provided by drug sponsors. For novel chemical entities, traditional sets of laboratory and field studies generally provide a basis for inferring post-approval safety and effectiveness. For drugs that have already been used in animal populations, such as unapproved drugs or new indications for approved drugs, there may be existing data from representative populations that are appropriate to support a regulatory decision. In the literature on evidence-based medicine, well-designed systematic reviews and meta-analyses that integrate information from population studies are considered to provide stronger evidence about treatment effects than individual randomized controlled trials. While methodologies and resources for conducting such reviews are now widely available, the use of these approaches in regulatory decision-making has some unique requirements to allow a conclusion: 'the drug is safe' and 'the drug is

effective.' However, systematic reviews and meta-analyses provide a rigorous framework for integrating evidence from published literature and other sources to strengthen regulatory decision-making.

Use of Literature to Make Regulatory Decisions: The FOLLTROPIN Example
Emily Smith, FDA/CVM

The FOLLTROPIN approval highlights different ways in which literature may be used as part of regulatory decision-making. The sponsor conducted a systematic review and meta-analysis to provide substantial evidence of effectiveness of FOLLTROPIN, instead of conducting a traditional field study, because of the availability of a large number of published and unpublished studies. The regulatory systematic review differed from a typical systematic review (e.g., Cochrane review) in the scope of the review question, the design of the study eligibility assessment, the evaluation and inclusion of sponsor-conducted studies, and the extent of data extraction and bias assessment for individual studies. The review question was defined with a very narrow scope that included a consideration for a particular formulation of drug product, used in dairy and beef heifers and cows, under specific conditions of use. The systematic review and meta-analysis allowed the Agency to conclude that FOLLTROPIN was effective for the proposed conditions of use, that the results are likely to be repeatable, and that valid inferences can be drawn to the target population. For the target animal safety evaluation, although a systematic review was not performed, literature was used to characterize the pharmacology and toxicology of FOLLTROPIN, justify an alternative approach to the traditional margin of safety study, and identify remaining gaps in the information necessary to complete the target animal safety evaluation for FOLLTROPIN.

A Retrospective Meta-Analysis on the Effectiveness of FOLLTROPIN-V for the Induction of Superovulation in Beef and Dairy Heifers and Cows
Steven Radecki, ASA; Junshan Qiu, FDA/CDER; Anna Nevius, FDA/CVM

Data from 21 research and clinical studies encompassing 50 treatment arms were included in the analysis. Variables included in the analysis were treatment arm, study, mean number of transferable embryos within a treatment arm, variation in the mean number of transferable embryos within a treatment and the number of records used in the estimation of the mean number of transferable embryos. Statistical assessments of bias included the variables percent missing, number of records used in the estimation of the mean, and the harmonized SEM. Publication bias and the heterogeneity among the treatment arms were also evaluated. The final statistical model based on these evaluations was a two-level random effects model (between study and between treatment arms) with intercept included as a random effect at the study level. A sensitivity analysis evaluating potential high-risk study characteristics was also conducted.

Effect size measurements were generally reported as means or least squares means. Standard effort of the mean was the measure of variation harmonization. The lower bound of the two-sided 95% confidence interval for the mean number of transferable embryos was greater than 1 for 36 of the 50 treatment arms. The estimated effect size based on the model was 4.5 transferable embryos (95% confidence interval: 3.5, 5.6). Given the lower bound of this interval was greater than the expected non-treated value of 1, it was concluded that treatment with FOLLTROPIN-V is effective in increasing the number of transferable embryos. The sensitivity analysis confirmed the validity of the chosen model.

Friday, September 18
8:30 a.m. – 9:45 a.m.

PS4A

Advancing Personalized Medicine Using Innovative Subgroup Identification Methods
Thurgood Marshall North

Organizer(s): *Qi Tang, AbbVie; Rong (Rachel) Chu, Agensys, Inc.; Yun Wang, FDA/CDER/OTS/OB/DB5; Anna Sun, FDA*
Chair(s): *Qi Tang, AbbVie*

Personalized medicine is the future of drug development. However, our limited understanding of human biology is a big hurdle for development of personalized medicines. To overcome this hurdle, several novel subgroup identification methods have been developed recently to better utilize the clinical data at hand to generate hypotheses about personalized treatments. There are many challenges faced by subgroup identification methods: variable selection bias, control of Type I error, multiplicity, predictive performance and confounding variables. Because of these challenges and the complication of subgroup identification methods themselves, it is almost impossible to know which method works well under which scenario. Thus, in practice, for a given data set, multiple methods need to be applied and the one with the best predictive performance chosen. The purpose of this session is to bring to the audience the most recent advances in subgroup identification methods and offer practical guidance for selecting appropriate methods.

Statistical Methods for Subgroup Identification in Personalized Medicine
James Chen, FDA/NCTR

Personalized medicine applies molecular technologies and statistical methods to identify genomic biomarkers in target patients for assigning more effective therapies and avoiding adverse events. Subgroup identification involves partitioning patients into subgroups defined by sets of biomarkers, where each subgroup corresponds to an optimal treatment. Subgroup identification for treatment selection consists of the three components: 1) biomarker identification; 2) subgroup selection; and 3) performance and clinical utility assessment. Biomarker identification involves developing statistical test procedures to identify a potential set of biomarkers to define patient subgroups. Subgroup selection is to develop a class prediction model to identify patient subgroups for treatment selection. Performance and clinical utility assessment evaluate 1) accuracy of classifiers and 2) power to detect treatment effect in the targeted subgroup. Statistical issues and challenges include experimental design, statistical models and tests to identify predictive biomarkers, classification model development to identify subgroups, classification of imbalanced subgroup sizes, and multiple testing.

The GUIDE Regression Tree Approach
Wei-Yin Loh, University of Wisconsin-Madison

In using a regression tree to identify subgroups with differential treatment effects, the first requirement is that the method be free of bias in the selection of variables to split the nodes. A second requirement is that the method be able to differentiate between the effects of prognostic and predictive variables. A third requirement is that it does so with good accuracy. This talk will discuss a recent extension of GUIDE that satisfies these requirements and compares its performance with that of other existing regression tree methods.

Exploratory Identification of Biomarker Signatures for Patient Subgroup Selection in Clinical Drug Development
Xin Huang, AbbVie

Mechanistic relationships between putative biomarkers, clinical baseline and related predictors versus clinical outcome (efficacy/safety) are usually unknown, and must be deduced empirically from experimental data. Such relationships enable the implementation of a personalized medicine strategy in clinical trials to help stratify patients in terms of disease progression, clinical response, treatment differentiation, etc. The relationship between some biomarkers and clinical baseline predictors versus clinical outcome are typically stepwise or nonlinear, often requiring complex models to develop the prognostic and predictive signatures. For the purpose of easier interpretation and implementation in the clinic, defining a multivariate biomarker signature in terms of thresholds on the biomarker combinations would be preferable. In this talk, we present some methods for developing such signatures in the context of continuous, binary and time-to-event endpoints. Further, to evaluate the future sample performance of the biomarker signature, we proposed the concept of predictive significance via cross-validation. Results from simulations and case-study illustration will also be provided.

PS4B

Modeling and Simulation at the FDA and in Industry: Collaboration Among Statisticians, Modelers, and Pharmacometricians
Thurgood Marshall South

Organizer(s): *Cristiana Mayer, J&J; Shiowjen Lee, FDA; Alan Hartford, AbbVie; Misook Park, FDA*
Chair(s): *Cristiana Mayer, J&J*

The increasing costs of drug development forces the industry and regulators to look at innovation in a more aggressive way. The role of modeling and simulation (M&S) has gained great momentum and enthusiastic interest from all stakeholders in recent years. This session will bring to the table the challenges and solutions for promoting model-based drug development and model-informed regulatory assessment to a new higher level. Quantification of risk, statistical and mathematical modeling, virtual trial simulation are all important tools in advancing the characterization of dose response profile, PK/PD relationship and benefit/risk ratio. This effort requires a more effective and intense collaboration between industry and regulators as well as among different groups within quantitative sciences, such as statisticians, modelers and pharmacometricians. All parties are motivated to improve and intensify the collaboration. The session will highlight a case study from a very recent M&S approach in a regulatory submission paired with the model-informed regulatory assessment leading to the approval. The perspectives, concerns and recommendations from the different groups involved will be emphasized.

Incorporation of Stochastic Engineering Models as Prior Information in Bayesian Medical Device Trials and Post-Market Surveillance
Tarek Haddad , Medtronic; Rajesh Nair, FDA/CDRH; Laura Thompson, FDA; Telba Irony, FDA; Adam Himes, Medtronic

Modern implantable medical devices have brought improved quality of life to many patients. Evaluation via clinical trial is often a necessary step in the process

of bringing a new product to market. In recent years, device manufacturers are increasingly using stochastic engineering models during the product development process. These models have the capability to simulate virtual patient outcomes. Incorporation of these models as prior knowledge in a Bayesian clinical trial design can provide benefits of decreased sample size and trial length while still controlling type I and type II error rates. This paper presents a straightforward method for augmenting a clinical trial using virtual patient data, where the number of virtual patients is based on the similarity between modeled and observed data. The use of this method is illustrated by a case study based on a model for cardiac lead fracture.

Model-Based Bridging of Dose-Regimens Supporting Drug Approval: A Case Study of Cross Functional Modeling Collaboration
José Pinheiro, J&J; Chyi-Hung Hsu, Janssen R&D

Modeling and simulation approaches can greatly improve the efficiency of clinical drug development, leading to faster and better quantitative decision-making, when properly applied. Different disciplines involved in drug development utilize modeling and simulation techniques as part of their methodological toolbox. This has led to some confusion, eventual misunderstanding and competition, but increasingly synergistic collaboration among disciplines. This presentation will discuss a case study based on a real fixed dose combination program in Type 2 diabetes mellitus in which modeling and simulation played a key role in the approval of the NDA. Collaboration among the different modeling disciplines involved was instrumental to the success of the project and will be illustrated and discussed in the talk.

Modeling and Simulation at the FDA and in Industry: Collaboration Among Statisticians, Modelers, and Pharmacometricians
Vikram Sinha, FDA

The talk will address the role of model-based approaches and model-informed regulatory assessments in drug development from a pharmacometric regulatory perspective with highlights on the collaboration between industry and the FDA.

PS4C

Messy Data Issues in Evaluation of Bioequivalence

Thurgood Marshall East

Organizer(s): *Wanjie Sun, FDA; Stella Grosser, FDA; Mark Shiyao Liu, Mylan Inc.; Charles DiLiberti, Montclair Bioequivalence Services, LLC*
Chair(s): *Julia Jingyu Luan, FDA/CDER*
Panelists: *Lisa LaVange, FDA; Yi Tsong, FDA; Shein Chong Chow, Duke; Pina D'Angelo, Novum Pharmaceutical Research Services*

In 2013, generic drugs accounted for 86% of the market share in U.S. However, statistical research in bioequivalence tests, especially how to handle complications such as missing data and messy data, has been lacking. Bioequivalence studies can be pharmacokinetic (PK) bioequivalence or clinical end-point bioequivalence studies. With the institution of the Generic Drug User Fee Act (GDUFA) in 2012, it becomes a pressing task to study the impact of data complications on statistical conclusions, and the robustness of the current statistical methods in face of these complications. Furthermore, robust statistical methods need to be proposed to handle complications in bioequivalence tests. This session will include two presentations followed by a panel discussion. Speakers and panelists from the agency, industry and academia will present and discuss issues and statistical approaches in the area of bioequivalence.

Missing Data and Noncompliance Data in Clinical End-Point Equivalence Studies
Wanjie Sun, FDA; Aotian Yang, GWU; Stella Grosser, FDA; Carol Kim, FDA

In clinical trials, patients may drop out for various reasons (e.g., lack of efficacy, treatment-related side effects, or factors unrelated to the trial) (Heyting et al 1992). Noncompliance also happens very often, such as poor compliance rate, out-of-window visits, and so on. Drop out and noncompliance can be balanced or imbalanced between treatment groups. Methods to handle missing data have been studied extensively in superiority trials for new drugs (Little and Rubin 2002, Molenberghs and Kenward 2007, Ibrahim JG and Molenberghs 2009, Siddiqui et al 2009, etc.). However, the impact of missing data and non-compliance data on bioequivalence trials for generic drugs, particularly, clinical end-point equivalence studies, has seldom been investigated. With the institution of the Generic Drug User Fee Act (GDUFA) in 2012, it becomes a pressing task to study the dropout pattern in the current ANDA submissions, to test the robustness of the current practice for handling missing and non-compliance data in clinical end-point equivalence studies, and to propose feasible sensitivity methods.

In this presentation, a meta-analysis will be discussed to evaluate the missing data/non-compliance patterns in the ANDA clinical end-point bioequivalence studies using topical drugs for treatment of acne vulgaris as an example. Simulation results will be presented to evaluate the bias and efficiency of the current practice under different missing data mechanisms (missing completely at random, missing at random, missing not at random) and different non-compliance mechanisms. Sensitivity analysis will be briefly discussed in bioequivalence tests.

Messy Data in PK Bioequivalence Studies
Charles Bon, Biostudy Solutions, LLC; Lindsey Katz, Biostudy Solutions, LLC

The basic statistical methods for evaluation of bioequivalence using pharmacokinetic data will be reviewed with emphasis on replicated design studies. The use of 3-period, two-treatment, reference-replicated crossover studies has become routine in bioequivalence testing of highly variable drugs. The four-period, two-treatment, fully replicated design is a close second, with some recent FDA guidance documents requiring this design. The FDA has given industry SAS code in some guidance documents for replicated design BE evaluations. Under ideal circumstances, this makes the statistical analyses fairly straightforward. Unfortunately, such ideal circumstances seldom occur in real studies and the biostatistician is often faced with messy data due to it being unbalanced, study conduct inconsistent with the planned design, or convergence problems. Several examples of messy PK data will be discussed, including situations involving subject dropouts and inestimable PK parameters, and having multiple, instead of single, dosing groups when scaled average BE evaluation was required. The discussion will cover the changes made to standard statistical analysis methods that enabled bioequivalence assessments in these messy data situations. The results of simulations evaluating the validity of these changes will also be presented.

PS4D

Recent Innovations in the Development and Application of Statistical Designs for Early-Phase Oncology Trials

Thurgood Marshall West

Organizer(s): *Yuan Ji, NorthShore University HealthSystem/The University of Chicago; Adam Hamm, Theorem Clinical Research, Inc.; Hui Zhang, FDA; Xian Zhou*
Chair(s): *Sue-Jane Wang, FDA*

In this session, we discuss innovative adaptive designs in early phase oncology, statistical aspects, and challenges in implementing the designs. Focus is placed on innovation in both new methodology development and applications in practice. We also evaluate and establish prerequisites for choosing designs that are most efficient in attaining the goals of the trial. Hay et al. (2014, *Nature Biotech.*) presented a miserable success rate in oncology drug development compared to non-oncology diseases, and early phase oncology trials exhibit the largest deficiencies. We assemble a panel of academic and industry speakers, chaired by an expert from FDA, to present and discuss some recent breakthroughs aimed at greatly improving the design and conduct of early-phase oncology studies. Novel statistical designs such as dose finding in two treatment cycles based on efficacy and toxicity will be presented as innovation in methodological research, and real-life experiences in applying the mTPI design (Ji and Wang, 2013, *JCO*) with comparison to other classical designs such as 3+3 and CRM (O'Quigley et al., 1990, *Biometrics*) will be shared. Discussions will focus around requirements for implementing the various methods to ensure efficiency and reliability. Theoretical comparisons between different analysis methods, including optimal sample size among the methods will be described.

Utility-Based Bayesian Adaptive Designs for Early-Phase Clinical Trials
Peter Thall, MD Anderson Cancer Center

When deciding how to treat their patients, physicians must consider risk-benefit trade-offs between possible good and bad clinical outcomes. When designing clinical trials to evaluate new treatments, a practical approach that reflects this common medical practice is to elicit numerical utilities from the physicians planning the trial that quantify the desirability of each possible clinical outcome a patient may experience. In this talk, I will discuss how this may be done to design Bayesian sequentially adaptive early phase trials. After some preliminary remarks on Bayesian statistics, three illustrations will be presented. These include early phase trials involving (1) dose-finding for radiation therapy of pediatric brain tumors; (2) optimizing sedative dose in premature infants who must be intubated to treat respiratory distress syndrome; and (3) constructing a utility surface for two event times to jointly optimize dose and schedule in stem cell transplantation.

Implementation of Innovative Adaptive Designs in Early-Phase Oncology Trials: From Theory to Practice
Inna Perevozskaya, Pfizer Inc.

The methodology development for first innovative dose-escalation designs dates back to a couple of decades ago. Today many methods are available for statisticians to use and their statistical efficiency over traditional 3+3 design is well documented by numerous publications. Despite these methodological advances, the general level of acceptance of innovative

adaptive designs in Phase 1 oncology remains low. The barrier is often not the methodology but rather cautious attitudes towards novel designs and implementation challenges.

In this presentation, we will review two case studies utilizing adaptive design in a phase 1 oncology trial: one implementing a continual re-assessment method (CRM) and the other one using modified toxicity probability interval (mTPI) method. A review of study design, implementation, and lessons learned will be given. We will focus on practical aspects and challenges of implementation for both of these methods.

Practical Considerations for Adaptive Dose-Finding in Phase I Oncology Studies Using Toxicity Probability Interval Approaches
Linda Sun, Merck; Christine Gause, Merck

In early-phase oncology dose finding studies, the goal is to identify the maximum tolerated dose that has the potential to be efficacious for a single agent or a combination of two or more agents. We discuss the toxicity probability interval (TPI) approach (Ji et al. 2007) and the calibration-free modified TPI approach (Ji et al. 2010), which are adaptive designs that allow for the development of decision rules for toxicity intervals using a Bayesian framework. These types of designs are appealing to non-statisticians because they are easy to implement and dosing decisions can be pre-specified in the protocol. The statistical and practical considerations for the TPI/mTPI including choice of target toxicity rate, evaluation window, and MTD estimation will be presented along with examples illustrating the methods for both single agent and combination trials.

PS4E

Meeting the Ebola Challenge

Lincoln 5

Organizer(s): *Estelle Russek-Cohen, FDA/CBER; Deepak Khatry, MedImmune; Dionne Price, FDA/CDER; James Lymp, Genentech*
Chair(s): *Deepak Khatry, MedImmune*

The Ebola outbreak in West Africa has created multiple challenges in terms of public health and an opportunity for statisticians to respond to a public health crisis. We have neither an approved vaccine for Ebola nor an FDA approved therapeutic for the treatment of Ebola. The response on the part of agencies within Health and Human Services, the World Health Organization and regulatory agencies within Africa along with biopharmaceutical companies has been unprecedented. We plan on bringing experts together that have weighed in on study designs, considering issues such as ethics of randomized trials, supply constraints, multiple candidate products and the need to get products to people rapidly. Some designs that have been proposed to date specifically in the context of Ebola have included platform trials for therapeutic products that allow multiple therapies to be tested in the same trial and for vaccines, several forms of cluster designs. Because of the declining epidemic, data from animals may be used to inform the evaluation of medical products for Ebola. The session will focus on both therapeutic trials and vaccine trials.

Statistical Challenges in Developing Immune Correlates to Support Licensure of Ebola Vaccines
Ivan Chan, Merck; Kenneth Liu, Merck

Given the current, unprecedented Ebola epidemic in West Africa that has caused more than 26,900 reported cases with more than 11,100 deaths [WHO Situation Report, 5/20/2015], a high priority has been placed on development of prophylactic vaccines. International partnerships among governments, WHO, and industry have enabled rapid clinical development of an rVSV Ebola vaccine with multiple phase I to phase III studies being conducted simultaneously to support licensure. An 'accelerated approval' pathway based on immunogenicity data is also being considered in case the ongoing efficacy trials become inconclusive due to the declining incidence of Ebola. Therefore, it is important to identify immune correlates that can be used to support the basis of licensure. In this talk, we will discuss some key statistical challenges in developing immune correlates for Ebola vaccine, including endpoint selection, success criteria, and the use of animal (e.g., non-human primate) challenge model. Examples and simulations will be used to illustrate the proposed methods.

The Ebola Medical Counter Measures Trial: A Flexible Randomized Clinical Trial for Evaluating Therapeutics for Ebola Disease
Lori Dodd, NIAID; Michael Proschan, NIAID

Ebola is a deadly disease, with an estimated overall mortality rate of about 50%. Because the mortality rate is high, there is pressure to give experimental treatments—any treatments—to Ebola sufferers even without proof of efficacy. Interpretation of the resulting hodgepodge of data from ad-hoc allocation of treatments is nearly impossible. A randomized clinical trial is desperately needed to confirm efficacy of new treatments. We describe a flexible design allowing incorporation of promising new agents and facilitating interpretation of results even if the trial stops early because the epidemic wanes, the treatment supply is interrupted, etc. We dub the approach "barely Bayesian" because our non-informative uniform priors for the mortality probabilities in each arm are quickly overwhelmed by actual

data. Focusing on posterior probabilities rather than type I error rate makes the conclusions depend only on results observed, not what action would have been taken had other results been observed. We will discuss this and other features of the design.

PS4F

Minimizing Bias in Medical Device Trials Through Study Design and Data Analysis

Lincoln 6

Organizer(s): *Laura Lu, FDA/CDRH; Theodore Lystig, Medtronic, Inc.; Chia-Wen Ko, FDA*
Chair(s): *Laura Lu, FDA/CDRH*

Randomized and well-controlled studies are golden standard in clinical trial practice. However, there are situations in device studies where randomization is impossible, difficult, or potentially inappropriate. For example, investigators may face an ethical dilemma in recommending a randomized study to subjects when they believe that the different interventions in the study are not equally safe and effective (i.e., they lack clinical equipoise). Also, due to the implantation and operation procedures of some devices, it is impossible to keep the patients, clinician or evaluator blinded/masked. Improper trial monitoring could also lead to the break of blinding. Lack of randomization and blinding randomization will potentially lead to selection and operational bias and could adversely impact the level of evidence provided by the study and the ability to rely on the data as valid. In this session, we will focus on the approaches in minimizing bias in device trials through adequate study design, monitoring and data analysis.

Good Practice of Objective Propensity Score Design for Premarket Nonrandomized Medical Device Studies: A Discussion with Examples

Yunling Xu, FDA/CDRH; Heng Li, FDA/CDRH/OSB; Vandana Mukhi, FDA/CDRH/OSB; Nelson Lu, FDA/CDRH; Lilly Yue, FDA/CDRH

Nonrandomized comparative studies have been playing an important role in the premarket evaluation of medical devices. For such studies, objective study design using propensity score methodology is a core component. In this presentation, good practice of objective propensity score design for nonrandomized comparative studies will be discussed and illustrated with examples in a regulatory setting.

Consideration of Trial Design Comparing RCT to Single-Arm Study When Abundant Patient-Level Historical Data Are Available

Peter Lam, Boston Scientific

Developing trial designs in randomizing patients undergoing percutaneous coronary intervention (PCI) into short term Dual Anti-Platelet Therapy (DAPT) duration versus long term DAPT duration are challenging and may not be ethical in light of the recent data release showing 30 months of DAPT was superior than 12 months of DAPT with respect to major adverse coronary endpoints at the expense of slightly higher major bleeding rate. However, there is a clinical need to demonstrate a shorter DAPT duration may be of benefit for patients who are at high risk for bleeding undergoing PCI with a novel drug eluting stent.

To minimize the number of patients who are high risk for bleeding exposed to the long DAPT duration, one design option is a RCT with an unequal allocation based on a Bayesian conditional borrowing strategy with multiple cutoffs incorporating the historical prior extracting the target patients from the DAPT database with the control arm. This would potentially decrease the sample size needed further in the control arm in the new trial.

Another design option is a single arm study design compared to the historical control using propensity score quintile approach extracting the target patients from the DAPT database. This eliminates the need to expose any patients who are at high risk for bleeding to the long DAPT duration. The two design options will be compared in terms of sample size requirement, operating characteristics, and operating challenges in recruiting patients.

Balance Reduction for Observational Studies Using Propensity Scores

Thomas Love, Case Western Reserve University

The design of an observational study can mimic the approach of a randomized clinical trial in many ways, barring, of course, the most important element - randomization. Propensity scores and related methods (like the prognostic score) have been used to help alleviate observable selection bias in comparative effectiveness studies for some time. We present a new approach to assessing the impact of propensity score analyses on selection bias reduction in observational studies that combines the use of dot plots with more sophisticated assessments of distributional similarity. So-called Love plots use familiar graphical forms to help identify potential problems with pre- and post-adjustment standardized differences (or similar metrics) in the means of observational study covariates where matching,

weighting or other approaches are planned to account for observable confounding by indication.

A potential weakness of this approach is that the choice of metric inherently focuses attention on specific summaries of the distribution of the covariates involved that may or may not be sufficient for the eventual analytic task. We propose a new plot, making use of several augmentations that may help to alleviate the problem that balance in standardized differences may be insufficient to declare the covariates sufficiently balanced to merit further regression-based analyses without heroic assumptions. We demonstrate the use of the resulting plots in the design of several analytic approaches for assessing the impact of a particular medical device (the Swan-Ganz catheter) on a variety of outcomes in the context of the SUPPORT study, for which fairly complete data are available to the public.

Friday, September 18
10:00 a.m. – 11:15 a.m.

PS5A

Recent Developments and Considerations for Personalized Medicine: Follow-On 'Me Too' Companion Diagnostic Devices

Thurgood Marshall North

Organizer(s): *Yuying Jin, FDA/CDRH; Laura Yee, FDA/CDRH; Alicia Toledano; Kuang-Lin He, Fujirebio Diagnostics, Inc.*
Chair(s): *Qin Li, FDA/CDRH*

It is a current trend that diagnostic testing is used to select patients for corresponding therapeutic products. This is a medical model that involves both therapeutic products and companion diagnostic devices. With the rapid development in the area of personalized medicine, and a number of FDA-approved companion diagnostics for use with specific corresponding therapeutic products, opportunities exist for device companies to develop follow-on companion diagnostic devices, which also are called "Me Too" companion diagnostic devices. These follow-on companion diagnostic devices seek the same therapeutic indication as a FDA-approved companion diagnostic. There are many challenges associated with the validation of these devices, including patient sample availability, the lack of therapeutic partner, etc. In the session, we will discuss the study design and statistical considerations for "Me Too" companion diagnostics from the perspective of FDA researchers, academia and industry.

Biomarkers Recent Developments and Considerations for Personalized Medicine: Follow-On 'Me Too' Companion Diagnostic Devices
Xiao-Hua Zhou, University of Washington; Yunqi Bu, University of Washington

Personalized medicine is gaining more attention in medical research and practice. A market-ready companion diagnostic assay (CDx) is used in personalized medicine for choosing the best treatment for an individual patient. In the ideal situation, the CDx is used for patient enrollment in device-drug pivotal clinical trial(s) so that Food and Drug Administration can ensure that appropriate clinical and analytical validation studies are planned and carried out for CDx.

Unfortunately, development of the CDx may lag behind the development of the drug, and consequently, it is unavailable during the drug pivotal clinical trial. Instead, a clinical trial assay (CTA) may be used to enroll patients in the trial. Thus when CDx is available, to estimate the drug efficacy in the CDx intended use population, a bridging study will be required to assess the agreement between CDx and CTA in order to bridge the clinical data from CTA to CDx. The main challenge we face in a bridging study is covariate imbalance between treatment arms for the subpopulation with both positive CDx and CTA. In this paper, we introduce an estimation method for estimating the effect of a drug in a bridging study with missing data under a causal inference framework.

Building Bridges for Companion Diagnostics
James Ranger-Moore, Roche; Crystal Schemp, Roche

In 2012, Ventana Medical Systems faced the challenge of bridging one of its immunohistochemistry (IHC) assays to an already FDA-approved fluorescent in-situ hybridization (FISH) assay for selecting assay-positive patients for treatment with a targeted therapy in lung cancer.

This bridging strategy was motivated by IHC's lower cost, ability to use bright-field microscopy (thereby preserving morphologic information), and wider availability. There was an ongoing phase III selection trial that provided the opportunity for comparing the IHC assay to the FISH assay.

An IHC scoring algorithm was developed to maximize concordance with FISH while maintaining potential for high reader precision through algorithm simplicity. The algorithm training occurred in independent samples prior to its use in the phase III trial.

The primary endpoints for the bridging study were positive percent agreement (PPA) and negative percent agreement (NPA) of the IHC assay with the FISH

assay. The secondary endpoint was progression-free survival (PFS), with the analysis focused on the hazard ratio (HR) between two treatment arms. Statistical analysis had to consider a number of limitations, including that (1) some cases were unavailable for IHC analysis; (2) results for IHC+ cases that were FISH- (or unevaluable) were not available (and also, being screen fails, had limited amounts of other clinical information available); (3) there were multiple plausible ways to define best and worst case scenarios where imputation or simulation occurred, and (4) the trial's inclusion/exclusion criteria were designed with FISH, not IHC, in mind.

The complex nature of analysis in this setting highlighted many of the challenges of performing a successful bridging study for a companion diagnostic device; the lessons learned during this undertaking are also presented.

Study Design and Statistical Considerations/ Challenges for 'Me Too' Companion Diagnostics
Meijuan Li, FDA

It is a current trend that diagnostic testing is used to select patients for corresponding therapeutic products that involves both therapeutic products and companion diagnostic devices. With the rapid development in the area of personalized medicine, and a number of FDA-approved companion diagnostics for use with specific corresponding therapeutic products, opportunities exit for device companies to develop follow-on companion diagnostic devices, which also are called "Me Too" companion diagnostic devices. These follow-on companion diagnostic devices seek the same therapeutic indication as a FDA-approved companion diagnostics. There are many challenges associated with the validation of these devices, including patient sample availability, the lack of therapeutic partner, etc. In this talk, we will discuss the study design and statistical considerations/challenges for the clinical validation of "Me Too" companion diagnostics.

PS5B

Innovative Designs and Advanced Statistical Methodologies for Rare Disease Clinical Trials
Thurgood Marshall South

Organizer(s): *Yeh-Fong Chen, FDA; Hope Knuckles, Abbott; Laura Johnson, FDA; Jeffrey Krischer, University of South Florida*
Chair(s): *Yeh-Fong Chen, FDA*

Even though prevalence of each rare disease is low, roughly 30 million Americans have been affected by one or more of the nearly 7,000 rare diseases. For most rare diseases, it can be challenging to conduct clinical trials with enough power to detect the treatment effect. To bring a breakthrough therapy to the market early, it is important to find efficient approaches to utilizing individual patient data (e.g., improved study design and sound statistical methods). Although it may be necessary to adjust the general standard in clinical trials for common diseases when it is applied to rare diseases, it is not clear how the general standard should be adjusted to ensure both the quality of good trials and the efficacy of approved drugs.

Many workshops have been run in recent years to accelerate the development of therapies for rare diseases. Researchers from the industry, academia and regulatory agencies are working diligently to develop innovative trial designs and statistical methodologies that can be applied to this area. Nevertheless, more research is needed for reaching a consensus. Emerging topics include the use of two or three staged enrichment designs to target specific type of patient populations and the use of historical controls to efficiently conduct trials that will reduce the number of subjects recruited and ease ethical considerations. For both cases, Bayesian approaches have been proposed; however, its usage in terms of applications is not widespread.

Three speakers from different organizations will present their successful work in the rare disease area. This session provides a platform for researchers working in this area to discuss the challenges they faced, share the lessons they learned, and offer possible solutions.

A Sequential Multiple Assignment Randomized Phase 2 Trial for Rare Diseases
Roy Tamura, University of South Florida

Clinical research in rare diseases is difficult for a number of reasons, including limited number of patients, and patient attitudes toward active and placebo therapy. Oftentimes, a number of available drugs are used for a rare disease with limited information about the efficacy of any of the drugs. In such situations, it is logical to first design a trial that would determine if any of the drugs shows promising efficacy. In cooperation with the Vasculitis Clinical Research Consortium under the NIH, we have recently developed a sequential multiple assignment randomized trial (SMART) for rare vasculitis diseases. The goal of the design is to determine the best out of a number of potential drugs and to compare the chosen drug to the best of the remaining drugs. In this talk, I will present the design, analysis, and operating characteristics of the SMART for three drugs when the number of subjects is limited.

Challenges Encountered in Conducting Rare Disease Clinical Trials
Min Min, FDA

Rare diseases literally mean diseases that are occurring in small patient populations. In extreme cases, the number of patients who can participate in a clinical trial can potentially be fewer than ten. At times it is not feasible to conduct a traditional, so-called, "adequate and well controlled" clinical trial with type I error control. In studying these ultra-rare diseases, two major challenges are encountered: 1) lack of sufficient study power to detect the treatment effect and 2) lack of comprehensive disease knowledge for identifying suitable endpoints for assessing drug effect.

To tackle these challenges, potential solutions can be: 1) exploring treatment effect in terms of patient profiles; 2) adopting innovative trial designs; and 3) comparing trial data with natural history data. In particular, regarding patient profiles, we can use them to examine patients' disease progress and improvement in regards to the timing of treatment interventions. In designing a trial for a rare disease it might be unethical to include a concurrent control. In this circumstance, a multi-arm trial that includes a well-defined natural history study may be considered. In including an historical control, it may useful to consider whether we should potentially use a Bayesian statistical approach.

In my presentation, I will share my FDA experience in evaluating rare disease clinical trials. Several cases with different types of challenges will be discussed. I will also offer the audience my recommendations.

Interim Futility Analysis Based on Linear Regression with Longitudinal Endpoint in a Rare Disease Indication
GQ Cai, GSK

There are many clinical trials where the primary endpoint is observed at a specific long-term follow-up time, while repeated measures of the same outcome are also taken at earlier visits. In such cases, it is possible that interim futility analyses will be planned such that the trials can be terminated early if the treatment does not induce any benefit to the patients. For such trials subjects only provide data on the primary endpoint once they have completed the long-term follow-up time, potentially eliminating a large proportion of the enrolled subjects from an interim analysis. We propose a more efficient interim analysis based on the slope of a linear regression, which incorporates all the data available at the interim analysis. This approach has the added advantage of providing a data-drive decision about the timing of the interim. The construction of interim futility rules and the timing of the interim analysis are discussed and the method is illustrated with an example involving a placebo-controlled comparison of longitudinal proteinuria measurements in a rare renal disease.

PS5C

Role of Statisticians in CDISC Data Standards: From Developers to Users

Thurgood Marshall East

Organizer(s): *Weiya Zhang, FDA; Deborah Bauer, Sanofi; Stephen Wilson, FDA/CDER/OTS/OB/DBIII; Peter Mesenbrink, Novartis*
Chair(s): *Deborah Bauer, Sanofi*
Panelists: *Stephen Wilson, FDA/CDER; Susan Kenny, Maximum Likelihood, Inc.; Chris Holland, Amgen; Deborah Bauer, Sanofi*

High-quality data is essential for clinical trial design, statistical inference and decision-making. Substantial efforts have been dedicated across industry, academia, and regulatory agencies to develop data standards aligned with the models defined by the Clinical Data Interchange Standards Consortium (CDISC). CDISC standards support the clinical and non-clinical research process from protocol design through data collection, data exchange, data management, data analysis and reporting. The Prescription Drug User Fee Act (PDUFA V) set goals for FDA to develop guidance for industry on the use of CDISC data standards for the electronic submission of study data in applications (see *www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm*). FDA has been actively working on guidance requiring study data in conformance to CDISC standards and developing distinct data standards for therapeutic areas using a public process that allows for stakeholder input through open standards development organization. Industry sponsors have also been working to develop therapeutic area data standards aligned with the guidelines developed by CDISC and the FDA.

With these standards in place and many under development, how are we as statisticians employing these standards? What are the advantages and challenges of implementing these standards? In this session, we will invite speakers from industry and regulatory to share their experience with developing and implementing CDISC data standards, and how these data standards facilitate regulatory reviews and clinical research.

The Good, the Bad, and the Ugly: The Pharmaceutical Industry Perspective with CDISC Data Standards
Peter Mesenbrink, Novartis

Awareness of the importance of end-to-end CDISC data standards has grown multifold over the past five years among statisticians in the pharmaceutical industry. However, it is not always clear to all the magnitude of the impact that can be made with high quality analytical decision making as part of CDISC data standards development and implementation. In an era with data transparency across Pharma, the development of disease-area CDISC data standards not only impacts the quality of the NDA/BLA review but also impacts comparative effectiveness research as part of pricing and reimbursement and exploratory data analysis and the planning of future clinical trials through meta-analysis. Experiences where development and implementation have been done well and suboptimally will be discussed and what needs to be done as part of collaboration between industry and FDA statisticians to ensure that we are able to maximize the use of standardized data in the future.

CDISC Data Standards: A Statistical Reviewer's Perspective
Weiya Zhang, FDA

The Prescription Drug User Fee Act (PDUFA V) set goals for FDA to develop guidance for industry on the use of CDISC data standards for the electronic submission of study data in applications. Many applicants are currently submitting CDISC compliant data sets in NDAs/BLAs. What are the advantages and challenges in using CDISC data sets for statistical review? How do the SDRG (Study Data Reviewer's Guide) and ADRG (Analysis Data Reviewer's Guide), developed in collaboration with PhUSE, help reviewers understand the relationships between the submitted data and the study report?

In this presentation, the speaker will share review experiences with standardized data and related documentation, describing: 1) current CDER guidance requiring the submission of study data in conformance with CDISC standards; 2) the statistical review of CDISC-based submissions; and 3) the use of the SDRG and ADRG to support review.

PS5D

ICH E9 R1 Defining the Estimand and Sensitivity Analysis

Thurgood Marshall West

Organizer(s): *Estelle Russek-Cohen, FDA/CBER; Brent Burger, PAREXEL International; Guoying Sun, FDA*
Chair(s): *Joan Buenconsejo, AstraZeneca*

ICH E9 is a key guidance on the subject of designing clinical trials and is recognized by multiple regulatory bodies. A new addendum is being written and it deals with defining the estimand and sensitivity analysis, preferably in advance of conducting a study. Motivation for the ICH addendum comes from the NRC report on missing data issued in 2010 but also a desire to improve the way we design clinical trials to properly assess treatment benefit. Issues such as how we deal with missing data will clearly be part of the discussions but so will sensitivity of the interpretation of trial outcomes to assumptions made at the time the study was planned.

The session will have two speakers who are members of the new ICH group, namely Tom Permutt (FDA) and Devan Mehrotra (Merck). It will be followed by a panel that includes our speakers and Dan Scharfstein (Johns Hopkins), Craig Mallickrodt (Eli Lilly), and Frank Bretz (Novartis).

Tackling Missing Data: Control-Based Quantile Imputation and Tipping Point Analysis
Devan Mehrotra, Merck

In a typical randomized two-arm (test, control) longitudinal clinical trial, the endpoint of interest is not observed for dropouts. The resulting missing data problem is commonly tackled by invoking a missing at random (MAR) assumption and proceeding with a mixed model repeated measures (MMRM) analysis. If the MAR assumption is incorrect, the estimated treatment effect is biased for the primary estimand of interest, the latter defined as the true between-treatment difference in endpoint means in the entire study population based on complete or partial adherence to assigned treatment in the absence of rescue medication. Published methods that attempt to decrease the bias are somewhat complicated and involve additional assumptions. We propose a simple solution in which the implicitly imputed mean for test-arm dropouts in the MMRM analysis is explicitly replaced with a given quantile (default: 50%) of the estimated endpoint distribution for the control arm. Systematically varying the level of the quantile leads to a spectrum of estimated treatment effects and corresponding *p*-values that can be used for (i) benchmarking the MMRM results,

and (ii) assessing robustness of the study conclusions via the tipping point concept. Three real data sets are used to illustrate the proposed methodology.

Practical Causal Estimands
Thomas Permutt, FDA

The causal inference literature is focused on model-building for large observational studies with many covariates. In contrast, confirmatory clinical trials are usually analyzed by fairly simple methods, specified in advance, with few covariates. I aim to give practical advice on what causal estimands can be estimated or approximated by simple, prespecified methods. I believe there are only three kinds:

1. The intent-to-treat effect using retrieved dropouts

2. Composite outcomes incorporating dropout as an outcome in itself

3. The effect in a subset defined by principal stratification, especially the total direct effect

PS5E

Using Historical Data in Clinical Trials: Synthesis of Truth with Uncertainty

Lincoln 5

Organizer(s): *Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals; Pandurang Kulkarni, Eli Lilly and Company; Margaret Gamalo, FDA/CDER; Kerri Schoedel, Altreos Research Partners, Inc.*
Chair(s): *Satrajit Roychoudhury, BDM Oncology, Novartis Pharmaceuticals*

In recent years, Bayesian design and analysis have generated extensive discussions in the literature of clinical trial. For Bayesian methods, decisions have to be made at the design stage regarding "prior belief." When good prior information exists, Bayesian approach may enable this information to be incorporated into the statistical analysis. There is an intrinsic interest of leveraging historical data into the prior information for efficient design. Clinical trials including historical control information are used in earlier phases of drug development (Neuenschwander et al., 2010; Trippa, Rosnerand Muller, 2012; French, Thomas and Wang, 2012; Hueber et al., 2012), occasionally in phase III trials (French et al., 2012), and also in special areas such as medical devices (FDA, 2010a; Campbell, 2011; Chen et al. 2011), orphan indications (Dupont and Van Wilder, 2011) and pediatric studies (Berry, 1989). Noninferiority trials also rely on historical information, and hence have similar characteristics as historical control trials (FDA, 2010b). "Enriching" control arm of current trial with information from historical trial(s) holds the promise of more efficient trial design. This allows trials with smaller size or with unequal randomization (places more subjects on the treatment arm in a study). This enriches the amount of information both on the efficacy or safety of the current novel treatment including important secondary endpoints. Borrowing historical information can further facilitate analysis of important subgroups. The advantages and disadvantages of this approach include increased power, decreased sample, and effects on type I error. There are many ways of borrowing from historical data. Generally all these methods act to "pull" or "shrink" estimates from the current control arm toward point estimates from the historical study(s). Moreover generally these methods have some parameters governing the borrowing, and can be set by the user to either borrow extensively or minimally. In all these methods guidance on setting these parameters is vital. But in practice these methods for borrowing historical information are not well understood in terms of benefits, effects, and regulatory ramifications.

This session will focus on different areas of incorporating historical information into clinical trials, along with their advantage and disadvantages. This session will feature four prominent participants (three speakers and one discussant) from industry, academia, and regulatory agency discussing borrowing information in different framework.

On the Use of Co-Data in Clinical Trials
Beat Neuenschwander, Novartis Pharma AG

Historical data are important for the design of a clinical trial. Yet, these data are rarely used in the analysis of the actual trial. While justifiable in certain situations, ignoring historical data can lead to less accurate inferences, and, therefore, suboptimal decisions. After a review of the main approaches to using historical data, the framework is extended to co-data, which comprise all relevant (historical and concurrent) trial-external data. These data can be used for the inference of the parameter in the actual trial via meta-analytic models. While the use of co-data in clinical trials is attractive, it is also ambitious. For example, avoiding undue weight of co-data (relative to actual trial data) is important, which can often be achieved by plausible assumptions about between-trial heterogeneity and by allowing for nonexchangeability among trial parameters. Two applications with co-data will be discussed: a phase III trial with interim decisions based on co-data; and, a phase I combination trial in Oncology, which takes advantage of co-data from completed and ongoing single-agent trials.

A Prospective Bayesian Adaptive Trial with Hierarchical Borrowing from a Prior Single Arm Study
Jason Connor, Berry Consultants; Kristine Broglio, Berry Consultants

We describe an adaptive randomized controlled Bayesian device trial with historical hierarchical borrowing for a cardiac device with both prospective efficacy and safety outcomes. This prospective two-arm trial incorporates an adaptive sample size plus adaptive borrowing from a one-armed trial of a similar device in Europe. Using a Bayesian hierarchical model, the design is structured so if U.S. data is similar to EU data, then more borrowing occurs, and less U.S. data is needed. If the U.S. and EU data are different, less borrowing will occur and the U.S. trial will run to a larger sample size, providing independent data to substantiate the device's safety and effectiveness in an American patient population. We illustrate the trial, its operating characteristics and show how the actual trial progressed, stopping early for success, and leading to FDA approval.

Incorporation of Prior Information in a Clinical Trial: A Reviewer's Perspective
Manuela Buzoianu, FDA/CDRH

When there is value recognized in the prior clinically relevant data, Bayesian methodology may be adopted in designing a new trial in which such initial data is used as an informative prior. A reason is that the new trial might have smaller sample size or duration by borrowing strength from previous studies. In this talk I will discuss some aspects of informative borrowing in the context of medical devices, focusing on issues to be considered when such Bayesian trial is submitted to support regulatory approval. In particular, I will present some design considerations for the assessment of the Bayesian modeling and the prior and discuss issues such as overly influential prior and divergence of the actual trial results from the prior.

PS5F

Analysis and Interpretation of Human Abuse Potential Study Data

Lincoln 6

Organizer(s): *Yahui Hsueh, FDA; Hope Knuckles, Abbott; Kerri Schoedel; Chen Ling, FDA*
Chair(s): *Wei Liu, FDA/CDER/OTS/OB/DBVI; Marta Sokolowska, Grunenthal USA*

Human abuse potential studies are randomized crossovers, which use multiple measures of subjective effects, administered at multiple timepoints, at multiple doses, over multiple periods of study. Their designs include sensitized subjects with enrichment for "responders," as well as highly sensitive endpoints ("canary in the coal mine" approach). These studies are often geared towards excluding false negatives (i.e., concluding no abuse liability, when it exists), usually at the expense of false positives results (concluding abuse liability when there isn't any). The use of multiple doses and nonlinear dose-responses (inverted U-shape) associated with drugs of abuse can present additional challenges in interpretation, as can the use of subjective measures, in the form of high variability and "distributional" violations. These factors can make analysis and interpretation of these data difficult for statisticians and non-statisticians alike. Are we satisfied with the "I'll know it when I see it" approach to evaluating abuse liability? The multiple endpoints and sometimes "exploratory" nature of these studies may lead us to succumb to the natural temptations of "cherry-picking." Are we better off using a confirmatory approach? How do we know we're using the right tests for our data? What are some new approaches? This session discusses the complexities associated with analysis and interpretation of abuse liability data, and opens a dialogue between statisticians and non-statisticians about addressing challenges inherent in these data.

This topic has been discussed at only a few prior meetings, only one of these was a statistical conference in 2011. In addition, currently, FDA is revising both the 2010 and 2013 FDA guidance documents (i.e., *Assessment of Abuse Potential of Drugs*, 2010, and *Abuse-Deterrent Opioids - Evaluation and Labeling*, 2013). Discussion of the best statistical approach to analyzing these data among statisticians from both the FDA and the pharmaceutical industry is extremely important. We also want to urge statisticians on both sides to get involved in conducting research in this important area.

Statistical Issues in Design and Analysis of Human Abuse Potential Studies
Chen Ling, FDA

The determination of whether a new molecular entity has abuse potential involves a variety of assessments. These can include safety studies, chemistry studies, receptor binding studies, pharmacokinetic studies, animal behavioral studies, human abuse potential studies, an evaluation of adverse events that occurring during clinical safety and efficacy studies, and evidence of abuse from epidemiological studies (if available).

The human abuse potential studies play an important role in the assessment of abuse potential of new drugs. These studies are typically randomized, double-blind, placebo- and positive-controlled crossover investigations in which there are usually at least four

treatments. These treatments consist of at least two doses of a test drug, at least one dose of a positive control drug (a drug with known abuse potential that is scheduled under the Controlled Substances Act), and placebo. Study subjects are healthy recreational drug users who have experience with drug class associated with the test drug. Such a study has multiple abuse potential measures, multiple primary endpoints, and multiple comparisons. In recent years, besides general human abuse potential studies, a lot of human abuse potential studies were conducted for assessing abuse deterrent opioids.

There are many statistical issues in human abuse potential studies. For example, misuse multiplicity adjustments for controlling type I error rate, underestimating the sample size of the study, unfair comparison between an intact extended release test drug and the same dose of a crushed immediate release positive control, etc.

In this presentation, I will discuss some of these issues and give my recommendations.

Defining the Liability in Abuse Liability: Practical Approaches to a Complex Analytical Problem
Naama Levy-Cooperman, Altreos Research Partners; Kerri Schoedel

Human abuse potential studies are randomized crossover studies, with numerous measures, endpoints and comparisons. These studies use enriched populations and multiple doses of investigational drug, comparator(s) and placebo. These study design factors, along with the often nonlinear dose-responses and non-normal distribution of subjective responses can create challenges in analysis and interpretation of study results. This presentation will introduce some of the key concepts in design of abuse potential studies, and discuss some of the inherent issues associated with analysis and interpretation of these data from a non-statistician's perspective, including potential underlying causes of these issues from a design perspective. The presentation will also offer some practical approaches to addressing some of these issues.

Statistics: The Real Abuse Liability
Susan Spruill, Applied Statistics and Consulting

What is the reason for conducting a human abuse potential (HAP) study? Are we interested is finding signals that indicate abuse potential, or are we interested in "proving" there is no signal? Regulated HAP studies are often geared towards excluding false negatives (i.e., concluding no abuse liability, when it exists), usually at the expense of false positives results (concluding abuse liability when there isn't any). The pharmacokinetic/pharmacodynamic relationships associated with drugs of abuse can present additional challenges in interpretation, as can the use of subjective measures, in the form of high variability and "distributional" violations. Are we using appropriate statistical approaches for interpreting these data? Do our interpretations map to the real world? Are we abusing statistics to get the answers we want as opposed to the answers we need? This talk will discuss the pro and cons associated with a select set of recommended study designs and analyses.

Friday, September 18
12:45 p.m. – 2:00 p.m.

PS6A

Statistical Considerations of Delayed Treatment Effects in Cancer Vaccine Trials
Thurgood Marshall North

Organizer(s): *Shenghui Tang, FDA; Zhenzhen Xu, FDA/CBER; Marc Buyse, IDDI Inc.; Jonathan Norton, MedImmune*
Chair(s): *Shenghui Tang, FDA*

In a relatively short period of time, therapeutic cancer vaccines have entered the landscape of cancer therapy. In contrast to the conventional chemotherapeutic drugs, these novel agents stimulate the patient's own immune response to combat cancer. This indirect mechanism-of-action for vaccines poses the possibility of a delayed onset of clinical effect, due to the time required to mount an effective immune response and the time for that response to be translated into an observable clinical effect. The conventional design and analysis methods based on log-rank test, however, often ignore this delayed effect and result in underestimated sample size with insufficient power, failing to detect the potential effects of the vaccines. More innovative statistical methodologies are needed to address the unique characteristics of therapeutic cancer vaccines in the design and analysis of such a trial.

This session will feature speakers from experts in industry, academia and regulatory arenas who will present their research on design and analysis of cancer vaccine trials. Three major topics will be addressed: (1) Sample size calculation considering the delayed treatment effects in cancer vaccine trials; (2) Proper analysis of cancer vaccine trials with delayed treatment effects; (3) Statistical challenges in cancer vaccine trials development from regulatory perspective.

Sample Size and Power Calculation in Therapeutic Cancer Vaccine Trials with Delayed Treatment Effect
Zhenzhen Xu, FDA/CBER

In a relatively short period of time, therapeutic cancer vaccines have entered the landscape of cancer therapy and several large vaccine trials have been conducted. In contrast to the conventional chemotherapeutic drugs, these novel agents stimulate the patient's own immune response to combat cancer. This indirect mechanism-of-action for vaccines poses the possibility of a delayed onset of clinical effect, due to the time required to mount an effective immune response and the time for that response to be translated into an observable clinical response. The conventional designs, however, often ignore this delayed effect and result in an underestimated sample size with insufficient power. In this talk, we propose an innovative approach for sample size and power calculation by incorporating the delayed treatment effect in the design and analysis of such trials. The properties of the proposed method are to be evaluated both empirically and theoretically.

Sample Size and Power of Survival Trials in Group Sequential Design with Delayed Treatment Effect
Jianliang Zhang, MedImmune; Erik Pulkstenis, MedImmune

In study designs for randomized clinical trials with a survival endpoint, the log-rank test is commonly used and the treatment effect is hypothesized with a proportional hazards alternative. Recently, treatment effects frequently seen in successful cancer immunotherapy trials have been manifested through a delayed effect pattern with a lag time, raising challenges to the use of conventional study design hypotheses. In particular, when a trial with interim analyses is designed using a group sequential method, the expected treatment effect from a log-rank test statistic varies across analysis times and differs from the parameter specified in the alternative hypothesis. In this paper, we present statistical analytical work that formulates a design including interim analyses with a survival endpoint under a delayed treatment effect alternative. Closed-form solutions are provided for calculating power and sample size over varying study/follow-up times for the group sequential, delayed treatment effect design. The analytical work is also presented graphically and a simulation is conducted for validation.

Power Calculation for Log-Rank Test Under a Nonproportional Hazards Model
Daowen Zhang, North Carolina State University

The log-rank test is the most powerful nonparametric test for detecting a proportional hazards alternative; and thus, is the most commonly used procedure for analyzing time-to-event data in clinical trials. When the log-rank is used for data analysis, the power calculation should also be based on the log-rank test (Schoenfeld, 1983 *Biometrics*). In some clinical trials, treatment may not manifest its effect right after patients receive the treatment. Therefore, the proportional hazards assumption may not hold. We derive formulas for the asymptotic power calculation for the log-rank test under this non-proportional hazards alternative. Simulation studies indicate that the formulas provide reasonable sample sizes for a variety of trial settings. An example will be used to illustrate our methods.

PS6B

Designing Bioequivalence Studies for the Evaluation of Generic Drugs: Addressing Challenges Arising from Different Sources of Variability

Thurgood Marshall South

Organizer(s): *Elena Rantou, FDA/CDER; Fairouz Makhlouf, FDA/CDER; CV Damaraju, Janssen R&D; Susan Huyck, Merck*
Chair(s): *Fairouz Makhlouf, FDA/CDER*

When designing bioequivalence (BE) studies for generic drugs, different issues arise as a result of the presence of high variability. Between-subject variability is often considerable in parallel, crossover or even simple paired-sample designs. Within-reference variability characterizes highly variable drugs and is used as a scaling factor for determining BE. The criteria for assessing BE are based upon the type and magnitude of observed variability, the study design and the viable sample sizes. Such choices are crucial as they influence the sensitivity of the test to meaningful differences (consumer risk) and affect the chance of rejecting good products (producer risk). Different cases of generic drug studies such as solid oral dosage, long-acting injectable and locally acting dosage forms will be discussed. For these forms, a variety of statistical techniques will be presented such as ANCOVA with a special covariate and a modified scaled-average bioequivalence criterion, when the data set is a paired-sample design. Finally, recently developed method using scaled criterion for assessing drug interchangeability will be proposed and a numerical study will demonstrate its use for generic products.

An ANCOVA Approach to Reducing the Residual Variance and Sample Sizes Required for Parallel Design Bioequivalence Studies
Charles DiLiberti, Montclair Bioequivalence Services, LLC

Parallel design bioequivalence (BE) studies, which are commonly used for long half-life drugs, often pose substantial challenges for sponsors in that they employ

no replication, and thus cannot utilize the popular reference-scaled average bioequivalence (RSABE) method to control sample size. Furthermore, their sample sizes are dictated by between-subject variance, which is often substantially larger than the within-subject variance that dictates the sample sizes of crossover design studies. As a result, parallel design studies may easily require hundreds of subjects to achieve reasonable power. While this problem may arise for some solid oral dosage forms, it is a common problem for long-acting injectable formulations. Under some circumstances, the apparent terminal elimination rate constant (kel) is strongly correlated with the primary pharmacokinetic (PK) parameters AUC and Cmax, and yet is independent of the treatment effect. Therefore, incorporating ln(kel) as a covariate in an analysis of covariance (ANCOVA) of the ln-transformed PK parameters (i.e., ln(AUC) and ln(Cmax)) provides an opportunity to dramatically reduce the residual variance and consequently, the sample size as well, for parallel design studies. Incorporating other, subject-related factors and covariates into the ANCOVA model can further reduce the residual variance and required sample size. The conditions under which this type of ANCOVA approach is valid, as well as practical considerations, such as how best to address missing covariate values, will be discussed and illustrated with simulations and actual case studies.

Assessing Bioequivalence of Locally Acting Generic Products: Statistical Controversies and Arising Issues
Elena Rantou, FDA/CDER

In an attempt to determine whether a generic product is bioequivalent to its reference listed drug (RLD), the comparison of the test and reference distributions of a pharmacokinetic parameter is necessary. Usually, the response variable is a logarithmically transformed bioequivalence metric like the area under the curve (AUC), or the maximum concentration (Cmax) and the statistical test consists of comparing the average responses from the test and reference distributions.

For locally acting dosage forms like creams, gels and ointments, the available data often follow a paired-sample design where a subject is exposed to both the test and the reference formulations and the appropriate criterion for determining bioequivalence, is the two one-sided (TOST) confidence interval. Such approaches, although theoretically correct, cannot be always applied as the special nature of the data introduces challenges like the limited number of available subjects and/ or their replicates, the unusually high within-subject and between-subject variability and the presence of outlying subjects.

Concerns arising from the use of a bioequivalence criterion are related to the sensitivity of the test for detecting only meaningful differences and at the same time, not rejecting good products. Additionally, statistical power is affected by various factors like the bioequivalence limit, the sample size, the within-reference variability and the choice of the regulatory constants. These concerns are going to be discussed in reference to data coming from an in vitro permeation test, using human skin. The data will be analyzed using regular average bioequivalence and different forms of reference-scaled average bioequivalence and the advantages and imposed risks of each approach will be discussed.

A New Proposed Scaled Criterion for Drug Interchangeability
Shein-Chung Chow, Duke University

Criteria for assessment of bioequivalence for generic drug products are reviewed. These criteria include criterion for average bioequivalence, criteria for population and individual bioequivalence, and a recent proposed scaled average bioequivalence (SABE) criterion for highly variable drug products. In addition, following similar idea of IBE and the development of SABE, a new criterion for assessment of drug interchangeability is proposed. A numerical study was conducted to illustrate the use of the proposed criterion. In addition to the assessment of drug interchangeability for generic drug products, the proposed criterion can be applied to the assessment of drug interchangeability for biosimilar products.

PS6c

Summarizing Case Studies to Learn and Improve Confirmatory Adaptive Trial Design and Implementation

Thurgood Marshall East

Organizer(s): *Weili He, Merck; Paul Gallo, Novartis; Xuefeng Li, FDA/CDRH; Min (Annie) Lin, FDA/CBER*
Chair(s): *Paul Gallo, Novartis; Xuefeng Li, FDA/CDRH*

In the past decade, there have been an increasing number of confirmatory phase III trials that utilized adaptive designs. However, the uptake and general use of adaptive trial designs still seems relatively low (estimated at approximately 20%, according to Tufts CSDD survey). Among the main reasons are likely the added complexity of these designs compared to traditional designs, perceived risks of regulatory concerns, and lack of consensus on best practices for planning, implementing, and documenting these trials. The speakers in this session will present several completed confirmatory adaptive design case studies collected by the DIA ADSWG Best Practices (BP) Subteam, describe challenges in the design or conduct of these trials, and solutions that were adopted, and will

provide general guidance on improvements that could address concerns that led to certain types of adaptive designs being characterized as "less well understood" in the 2010 FDA draft AD guidance. Invited speakers and panelists from the DIA ADSWG BP Subteam and from FDA, industry, and academia will share their views on challenges and solutions for some key issues in the design or implementation of confirmatory adaptive design trials.

Addressing Challenges and Opportunities of 'Less Well-Understood' Adaptive Designs
Weili He, Merck; Paul Gallo, Novartis

The draft adaptive design guidance released by FDA included an overview of adaptive study designs that were termed as less -well understood. There was relatively little regulatory experience with these studies and their properties were not fully understood at the time of draft guidance release. To promote greater use of adaptive designs, especially those that were categorized as less well understood, the DIA ADSWG Best Practice Subteam has worked on describing and characterizing these less-well understood trials, identifying challenges related to these trials, and making suggestions on design or study conduct improvements. This talk will summarize the work from the subteam.

Case Studies of Less Well-Understood Adaptive Designs
Eva Miller, inVentiv Health Clinical

As we learned from the 2012 AD survey by Morgan et al and surveys undertaken by CBER and CDHR, adaptive designs have not grown to be as large a proportion of total clinical trials undertaken as advocates would have hoped. The best way to promote the uptake of adaptive designs is through case studies and lessons learned. The DIA ADSWG Best Practice Subteam collected a number of case studies and summarized the challenges and lessons learned from them. In this talk, we will describe these case studies and present some key challenges related to study design or conduct of these trials. We will also make suggestions for improvement based on the lessons learned from these case studies.

Panel Discussion
Sue-Jane Wang, FDA; Greg Campbell, formerly of FDA/CDRH; Boguang Zhen, FDA/CBER; Jerry Schindler, Merck; Eva Miller, inVentiv Health Clinical; Weili He, Merck

The panel will discuss the work to date by the DIA ADSWG Best Practice Subteam and provide guidance and thought on future work.

PS6D

Incorporating Patient Perspectives in Medical Product Life Cycle

Thurgood Marshall West

Organizer(s): *Bennett Levitan, Janssen R&D; Scott Braithwaite, New York University School; Telba Irony, FDA; Martin Ho, FDA/CDRH*

On November 4, 2014, the FDA published a *Federal Register* notice (FRN) to solicit input from stakeholders on strategies to obtain the views of patients in development of medical product and ways to account for patients' input in the regulatory process under the Food and Drug Administration Safety and Innovation Act. This FRN reflects that the principle of patient-centered health care has been widely accepted in the U.S. This session will describe how various types of quantitative patient preferences data can be elicited, assessed, and applied at different stages of medical product life cycles. In particular, the session will focus on how the patient preference data can potentially be incorporated into FDA's regulatory decision-making process in the benefit-risk assessment context. Experts from the FDA, the industry, and the academia will shed light on this emerging and important regulatory science research area.

Incorporating Patient Preferences into Regulatory Decision-Making
Telba Irony, FDA

In March 2012, the Center for Devices and Radiological Health (CDRH) at the FDA issued a guidance document listing the factors considered when the center makes benefit-risk determinations for approval of medical devices. A groundbreaking factor described in the guidance is "Patient Tolerance for Risk and Perspective on Benefit." The center recognizes that considering patient preferences is essential because only patients live with their medical conditions and consequences of the choices they make for their own care. Until now, this factor has not been formally considered in the regulatory setting.

To explore the use of patient preference evidence into regulatory benefit-risk determinations, CDRH sponsored a survey to elicit obese patients' preferences in choosing weight-loss devices.

In this presentation, we will describe the weight-loss device survey and present the survey results, which have been used to develop a powerful decision-aid tool for regulatory reviewers. The tool provides estimates of patients' benefit-risk trade-off preferences and also stratifies patients according to their risk-tolerance. We will conclude the presentation by sharing experiences in using patient preferences in regulatory process and talking about the draft guidance document on patient preference information—submission, review, and inclusion in device labeling.

Patient-Focused Benefit-Risk Assessment
Bennett Levitan, Janssen R&D

A growing number of industry, health authority and patient-driven initiatives have been paving the way towards engaging patients in the medical treatment development and regulatory review. A major focus has been on assessing the relative importance, or preferences, that patients place on the benefits, harms and other aspects of treatment. This presentation will introduce key elements of one such initiative, the Medical Device Innovation Consortium's Patient-Centered Benefit-Risk Framework, including (i) the nature of preference sensitive decisions; (ii) the type of regulatory situations where patient preference information is potentially valuable; (iii) the utility of different approaches to preference assessment at different points in the development life cycles; and (iv) the notion of approval for a subgroup of patents defined by benefit-risk preferences. We will show examples of how preference information can be used to inform development strategy and support a transparent and defensible benefit-risk assessment.

What Research Is Necessary to Determine Whether Particular Measures to Incorporate Patient Preferences into the Medical Product Life Cycle Leads to Decisions That Are More Preference-Concordant?
Scott Braithwaite, New York University School

Preference-concordant decisions (e.g., choosing the option that maximizes the expected value of a decision given a particular patient's preferences and valuations of possible consequences) is an increasingly important goal in patient-centered health care. In pursuit of this goal, regulators are considering measures (e.g., modifications to the regulatory processes and/or product labeling) with the intent of advancing preference-concordant decision making in situations where the decision of whether to use the product is preference-sensitive (e.g., those in which a there are multiple diagnostic or treatment options and the decision which option to pursue depends upon the particular preferences of the decision-makers). However, it will be necessary to study whether any regulatory measures undertaken indeed advance this goal and/or whether they cause unintended consequences. We will discuss a potential research agenda to determine whether regulatory measures result in more preference-concordant decisions as well as their impact on overall benefits and harms.

PS6E

Common Statistical Issues FDA Encounters
Lincoln 5

Organizer(s): *Heng Li, FDA/CDRH/OSB; Vandana Mukhi, FDA/CDRH/OSB; Brent Burger, PAREXEL International; Adam Hamm, Theorem Clinical Research, Inc.*
Chair(s): *Heng Li, FDA/CDRH/OSB*

FDA statisticians routinely review investigational plans for pivotal clinical trials. Of the issues that may arise, some are quite common. In this session, we discuss such common issues. We hope this discussion would improve the quality of submission and consistency of the review.

Statistical Issues in Regulatory Reviews of CBER Products
Shiowjen Lee, FDA

Center for Biologics Evaluation and Research (CBER) regulates blood products (both derivatives/components and devices maintaining safe blood supply), preventive vaccines (e.g., flu vaccines and childhood vaccines), tissue, cellular and gene therapies and therapeutic cancer vaccines. CBER products are different from those regulated by CDER with respect to working mechanism and manufacturing processes. These factors subsequently contribute challenges to the study design and statistical inferences of clinical trial data. Although the statistical principles may seem to be similar to those of products that are regulated by other FDA centers, there are unique statistical issues for CBER products. The objective of this presentation is to share experiences of statistical issues commonly encountered in CBER reviews. Recommendations and considerations will be provided to aid sponsors in their preparation of submissions.

Multiplicity and Type I Error Control
Laura L. Fernandes, FDA/CDER/OTS/OB/DBV

Clinical trials in oncology often have key primary endpoints like progression free survival (PFS), overall survival (OS) that aim to capture the efficacy of the study drug. These endpoints are usually tested using a hierarchical testing procedure so as to control the overall type I error. Given that OS is captured over a longer duration, studies are usually powered for both PFS and OS as primary endpoints or with OS as the first key endpoint to be tested after PFS. Since PFS is used as a surrogate endpoint for OS, the OS interim analyses (IA) are timed to coincide with the final PFS analyses while the final OS analysis is to be done when the required OS events are achieved.

In recent years, there has been an emphasis on including patient reported outcomes (PROs) as key endpoints as opposed to utilizing them for exploratory analyses. These endpoints could focus on a particular aspect of the instrument, for example time-to-deterioration of cough, and would have to be included in the hierarchy of the overall testing procedure. The timing for the analyses (IA versus Final) and the placement of these endpoints in the overall hierarchy is crucial so as to ensure success on maximum number of endpoints.

This talk will focus on the challenges faced when reviewing such applications that have many key endpoints, using various graphical methods for the multiplicity and in addition having multiple analyses at different time points (IA1, IA2, Final analyses).

Statistical Study Design Considerations for Medical Device Clinical Studies: From an FDA Reviewer's Perspective
Xu Yan, FDA/CDRH; Heng Li, FDA/CDRH/OSB; Vandana Mukhi, FDA/CDRH/OSB

In this talk, study design challenges frequently encountered in the planning stages of medical device clinical studies will be discussed. Issues in three types of study designs will be discussed from a statistical reviewer's perspective: randomized controlled trials, nonrandomized comparative studies, and single arm studies with performance goals. It provides a framework to aid statisticians and preparers of pre-market submissions in deciding what information to include (and not to include) in the statistical sections

PS6F

Missing Data in Diagnostic Device Studies: Methods and Case Studies

Lincoln 6

Organizer(s): *Bipasa Biswas, FDA; Xuan Ye, FDA; Vicki Petrides, Abbott; Kristen Meier, Illumina, Inc.*
Chair(s): *Bipasa Biswas, FDA*

Missing data are a prevailing problem for both therapeutic and diagnostic medical device studies. While missing data may be minimized by appropriate design and conduct of a clinical trial, it is still inevitable in such trials. For handling missing data, it is generally recommended to conduct sensitivity analyses to assess the robustness of the analysis results. Yet there are differences regarding issues and handling of missing data for therapeutic and diagnostic medical devices due to differences in study designs. Ignoring missing results while reporting diagnostic performance can be misleading. This session will focus on various types of missing data in diagnostic device studies. Practical issues and methodologies for handling missing data under different scenarios will be presented and discussed. Case studies using real clinical trial data will be presented to illustrate the problem.

Missing Data in Diagnostic Device Studies: Methods and Case Studies
Xiao-Hua Zhou, University of Washington

The accuracy of a diagnostic test can be measured by its sensitivity, specificity, positive predictive and negative predictive values. More generally, a receiver operating characteristic (ROC) curve may be used to represent the accuracy of a diagnostic test.

To calculate these measures, we need to determine the disease status of each patient in the sample. The procedure that establishes the patient's disease status is referred to as a gold standard. However, for some studies, only a subset of the patients with diagnostic test results are chosen to receive the gold standard assessment. If the study population consists of only verified cases, the estimated accuracy of the diagnostic test may be biased. This type of bias is called verification bias. In this talk, I will discuss how to correct for verification bias in estimation of the ROC curves and covariate specific ROC curves and their areas. I will illustrate verification bias correction methods with a data set from a large clinical study on Alzheimer's disease.

A Case Study of Handling Missing Data in Diagnostic Device Studies
Yuqing Tang, FDA

When conducting clinical trials, every effort should be made to achieve complete capture of all data. In practice, however, some missing values can arise for various reasons in a clinical trial, especially for diagnostic devices. For handling missing data, there are different methods depending on the underlying missing mechanism. From regulatory perspective, the primary method of handling missing data should be justified based on the expected data mechanism. To ensure that the diagnostic study results are not driven by way of handling missing data, alternate methods for imputing missing data based on different assumptions are recommended to evaluate the robustness of the study results. In this presentation, the speaker will share the experience of handling missing data in in-vitro diagnostic device. Different methods are used for handling missing data to assess the robustness of the device performance. The interpretation of corresponding analysis results will be discussed.

Missing Data in IVD Studies
Hope Knuckles, Abbott

Data may be missing in IVD clinical studies due to many different reasons. Real-world examples of missing data will be discussed. Two case studies will be pre-

sented in which data are missing, and reasons will be provided as to why the data are missing. Two methods of handling the missing data will be compared.

Friday, September 18
2:15 p.m. – 3:30 p.m.

PS7A

Platform Trials and Master Protocols: New Adaptive Designs Advancing Personalized Medicine

Thurgood Marshall North

Organizer(s): *Ohad Amit, GSK; Cristiana Mayer, J&J; Rajeshwari Shridhara, FDA; Lei Nie, FDA/CDER*
Chair(s): *Teri Ashton, GSK*

In the race toward drug development innovation, adaptive designs have played a major and increasing role in the last decade. The concept of platform clinical trials has come to the surface and is gaining incremental acceptance among pharmaceutical companies for advancing personalized medicine and the study of challenging and/or rare diseases. In 2010, the I-SPY2 trial was launched as one of the first platform trials. I-SPY2 is an innovative collaboration across five pharmaceutical companies in a phase II breast cancer trial. This collaborative approach, which aims to accelerate the identification of the most promising compound for a given population or the most suitable population for a given experimental drug, significantly reduces the cost, time and sample size in drug development. The same platform protocol can investigate multiple drugs and regimens for multiple and diverse subgroups of patients. In rare or difficult to enroll populations, platform trials offer the opportunity to steer important new therapies into a standing clinical trials infrastructure potentially expediting the availability of highly needed new therapies. This session will provide the audience with an overview of the statistical aspects defining platform clinical trials with emphasis on the design novelties and implementation efficiencies. Examples across a broad range of areas will be provided to highlight statistical innovations, regulatory perspective, and implementation efficiencies.

Statistical Considerations in Developing Master Protocols in Oncology

Lijun Zhang, FDA; Shenghui Tang, FDA

A master protocol can include multiple diseases, multiple treatments, or multiple molecular markers. It could be an umbrella trial, a cloud trial, or basket trial. A study with a master protocol can screen patient for multiple biomarkers and could lower screen failure rate effectively. In 2014, Lung-MAP, the lung Master Protocol, was launched and it has four sub-studies that include four novel agents targeting either PI3K, CDK4/6, FGFR, or c-MET for the 2nd-line treatment (those who have progressed on platinum-based doublet chemotherapy) of advanced squamous NSCLC. Patients with none of these abnormalities will be enrolled onto the anti-PDL1 study. Each sub-study includes a phase II component and the primary objective of the phase II component is to evaluate if there is sufficient evidence to continue to the phase III component by comparing progression-free survival. Statistical issues in developing master protocols will be discussed, which include sub-study assignment for patients who has more than one biomarker, estimation of treatment effect under different sub-study assignments when biomarker prognostic and/or predictive effect sizes are different, etc.

Utilizing Patient Data to Guide Treatment: All Patients Are Not the Same!

J. Kyle Wathen, Janssen R&D

When a patient visits their physician and learns they have a disease they often assume that any medical guidance they receive is personalized. However, in many diseases, such as cancer, it is not uncommon that a patient enrolls into a clinical trial where their treatment is selected by the flip of a coin, ignoring almost all information about them. Recent changes in the approach to clinical trials are attempting to "personalize" patient treatments. In this talk, I will describe a adaptive platform trial where patients are initially screened and observed to determine which of the experimental treatment they are most likely to respond to. When the patient enters the study, only treatments that are likely to benefit the patient will be considered. At any point during the study, treatments may be removed because it is unlikely that they provide benefit and new drugs may be added to target a specific set of patients.

Platform Trials: Statistical Efficiencies and Practical Examples

Scott Berry, Berry Consultants

Platform trials, in which multiple treatments, or multiple factorial treatments, are explored in a single protocol create huge statistical efficiencies. These efficiencies include having increased power, for far smaller sample sizes relative to non-platform trials. The idea isn't one of theoretical interest, in that there are many examples of emerging platform trials. These range from oncology, infectious disease, dementia, and even embedded trials within a learning health care system. In this talk, I will outline some of the statistical efficiencies, common statistical aspects of platform trials, as well as providing some examples of platform trials from a wide range of therapeutic areas.

PS7B

Poolable or Non-Poolable: Challenges and Solutions

Thurgood Marshall South

Organizer(s): *Ying Yang, FDA/CDRH; Yu Zhao, FDA/CDRH; Minjung Yoon.; Ying Yan, Helsinn*
Chair(s): *Yunling Xu, FDA/CDRH*

Almost all pivotal clinical studies are conducted in multiple centers and/or regions. Patients from different study centers, country/regions, may have different effectiveness and/or safety profiles. The differences may be due to different patient baseline characteristics, surgeon or physician's knowledge and skill level, patient care, etc. It is always a concern whether the data from different centers or regions can be pooled together. The estimated treatment effect may be biased and misleading if the heterogeneity among study sites, regions are ignored. In this session, we will discuss the approaches that are used to assess the inconsistency among study centers and/or regions. We will also discuss statistical methods used to estimate the treatment effect while adjusting for the aforementioned heterogeneity.

Data Pooling in Clinical Studies: Statistical Model and Methods as Well as Our Experience

Shun Zhenming, Sanofi-Aventis

This presentation will be composed of four parts:

1. Discussion on statistical hypothesis and models for demonstrating evidence in safety and efficacy where integrated or pooled evaluation can be considered. Our conclusion is that a pooling strategy could be more valid for safety evaluation; a pooled evaluation could be meaningful for efficacy if the statistical hypothesis is formulated correctly (Shun et al, 2005, *Stat in Med*).

2. A brief review of the existing methods for meta-analysis

3. Two real-life examples: An example (Wei-Xiang Qi et al: *Clinical Drug Investigation*, 2014, and *Eur J Clin Pharmacol*, 2013) based on the safety data from a cancer drug and another one (Angela Webster et al, 2010) based on the efficacy data from a drug of induction therapy for prophylaxis against acute rejection in kidney transplant will be presented together with our evaluation on the validity and interpretation of the outcome. The issues of the pooled data summary in the example are selection bias in study selection, the nature of the studies that provided the source data, and interpretation of the findings.

4. Conclusion: We summarize our conclusion and recommendation.

Poolability Analyses in Medical Device Trials: A Reviewer's Perspective

Yu Zhao, FDA/CDRH; Ying Yang, FDA/CDRH

Almost all pivotal medical device trials are conducted in multiple centers and/or regions. The device effectiveness and/or safety profiles might be different across centers/regions. Therefore, directly pooling the data together without evaluation of data poolability might lead to biased treatment effect estimate. In this presentation, we will discuss from a reviewer's perspective the challenges and possible approaches of evaluation of the site/region poolability and briefly touch on what to do if the data turn out to be non-poolable.

Case Study: Assessing Poolability in a Large Randomized Study on Dual-Antiplatelet Therapy

Joe Massaro, Boston University

A recently completed study assessed the effect of long-term dual-antiplatelet therapy (DAPT) on incident cardiovascular events in patients receiving drug-eluting stents. More than 9,000 patients were given open-label DAPT for 12 months, after which they were randomized to one of two study treatment groups: an additional 18 months of DAPT or 18 months of Placebo. Primary endpoints were stent thrombosis and a composite endpoint of major adverse cardiovascular and cerebrovascular events (MACCE; composite of death, stroke, myocardial infarction). The study was designed in response to a request from the United States Food and Drug Administration (U.S. FDA) to manufacturers of coronary stents, and was conducted through a public-private collaboration involving the U.S. FDA, eight funding major stent and pharmaceutical manufacturers, and Harvard Clinical Research Institute (HCRI). Subjects were enrolled into the trial by HCRI and from four post-marketing stent surveillance studies sponsored by stent manufacturers. Within each of these five studies, patients were enrolled from up to four regions (North America, Europe, Australia, New Zealand). Thus, both poolability of regions and poolability across the studies were of interest. Here we define poolability as consistency of treatment effect across regions and across studies. We will discuss our assessments of whether poolability existed, and how and why we adjusted for potential differences in baseline characteristics across studies and across regions in our poolability assessments, through the use of propensity scores.

PS7C

Emerging Topics in Benefit-Risk Assessment

Thurgood Marshall East

Organizer(s): *Weili He, Merck; Qi Jiang, Amgen; John Scott, FDA/CBER; Xuefeng Li, FDA/CDRH*
Chair(s): *Weili He, Merck; Xuefeng Li, FDA/CDRH*

In February 2013, FDA released a draft PDUFA V implementation plan on structured approach to benefit-risk assessment in drug regulatory decision-making. This document joined several other benefit-risk guidance/recommendations, including ones from FDA CDRH, PROTECT, EMA, and ISPOR, which have been widely reviewed and discussed in recent years. This body of BR guidance and recommendations is very timely and consistent with the long understanding that the benefits of a medical product can only be understood in the context of the risks or harms associated with that product, and vice-versa. The Quantitative Sciences in the Pharmaceutical Industry (QSPI) Benefit-Risk Working Group (BRWG), formed in early 2013, has been actively pursuing several emerging topics in BR assessment, including identification and evaluation of uncertainties in BR assessment, commonly used graphics in BR assessment in clinical development, identification of different data sources in BR assessment, and issues to consider for BR assessment in subgroups. This session will present the most current work from this working group. In addition, a panel of experts in BR assessment from FDA, industry, and academia will share their views on the current regulatory environment, emerging issues in BR assessment, and next steps and future directions.

Panel Discussion

John Scott, FDA/CBER; Telba Irony, FDA; Ellis Unger, FDA/CDER; Qi Jiang, Amgen

The panel will discuss the work to date by the QSPI Benefit-Risk Working Group and provide guidance and thought on future work.

Emerging Topics in Benefit-Risk Assessment

Qi Jiang, Amgen; Weili He, Merck; John Scott, FDA/CBER

Much emphasis has been placed on the structured benefit-risk assessment globally. The Quantitative Sciences in the Pharmaceutical Industry (QSPI) Benefit-Risk Working Group (BRWG) has been actively pursuing several emerging topics in BR assessment. This presentation will cover the most current work from this working group, with emphasis on a few important emerging topics in BR assessment, including recommendation of the methods, identification and evaluation of uncertainties, graphics, and identification of different data sources.

PS7D

Bayesian Assessment of Benefit-Risk Balance in Drug Development

Thurgood Marshall West

Organizer(s): *Maria Costa, GSK; Yueqin Zhao, FDA; Carl DiCasoli, Bayer Healthcare Pharmaceuticals; Min Min, FDA*
Chair(s): *Yueqin Zhao, FDA*

To gain regulatory approval, a new medicine must demonstrate that its benefits outweigh any potential risks. Over the past several years there has been a growing recognition amongst sponsors and regulators for the need of a more structured and consistent approach in assessing the benefit-risk balance of new therapies. The Bayesian inference framework and philosophy offers a tool for learning and updating one's beliefs about particular parameters of interest. This aspect of the Bayesian philosophy is especially attractive in the context of benefit-risk assessment, as existing information can be formally incorporated into the analysis of any emerging data. In addition, posterior probabilities offer a simple and clear device with which one may convey the benefit-risk balance to a non-statistical audience. This session will feature three talks showcasing the added benefit of including a formal assessment of the benefit-risk balance using Bayesian inference from different perspectives: sponsor, regulatory, and academia. The objective is to understand the potential for these methods to provide greater clarity of the benefit-risk balance to regulators, and what the state of the art is regarding statistical methodology.

What Role Should Formal Risk-Benefit Decision-Making Play in the Regulation of Medicines?

Deborah Ashby, Imperial College London

The regulation of medicine requires evidence of the efficacy and safety of medicines, and methods are well developed to deal with the latter and to a lesser extent the former. However, until recently, assessment of risk- benefit especially in relation to alternatives has been entirely informal. There is now growing interest in the possibilities of more formal approaches to risk-benefit decision-making. In this talk, we review the basis of drug regulation, the statistical basis for decision-making under uncertainty, current initiatives in the area, and discuss possible approaches that could enhance the quality of regulatory decision-making, drawing on experience of the IMI PROTECT project.

Bayesian Approach to Personalized Benefit-Risk Assessment with Application to Clinical Trial Data
Ram Tiwari, FDA

Benefit-risk assessment is critical in evaluating the effectiveness of a new treatment over the existing ones. Some benefit-risk measures depend on the probabilities of benefit-risk categories in which the subject-level benefit and risk outcomes are characterized. The existing benefit-risk methods for analyzing the categorical data depend only on the frequencies of mutually exclusive and collectively exhaustive categories that the subjects fall in, and thus ignore the subject-level differences. We propose a Bayesian method for analyzing the subject-level categorical data with multiple visits. A generalized linear model is used to model the subject-level response probability of each category, with respect to a "reference" category, assuming a logit model with subject-level category effects and multiple visit effects. Dirichlet process is used as a prior for the subject-level category effects to catch the similarity among the subject responses. We develop an efficient Markov chain Monte Carlo algorithm for implementing the proposed method, and illustrate the estimation of individual benefit-risk profiles through a simulation study. A clinical trial data is analyzed using the proposed method to assess the subject-level or personalized benefit-risk in each arm, and to evaluate the aggregated benefit-risk difference between the treatments at different visits.

A Novel Methodological Approach Which Allows Structured Benefit-Risk Assessments to Incorporate Both Uncertainty and Correlations Between Endpoints and Weights
Ian Hirsch, AstraZeneca

Structured benefit-risk assessments are being implemented across most internal AstraZeneca projects through a core benefit-risk implementation team via the BRAT framework. This has facilitated a systematic evaluation and allowed for a more transparent and standardized way of presenting a single benefit-risk assessment for each compound facilitated using an internally developed BRAT tool. Most assessments so far have been qualitative in nature using discrete comparisons between key benefits, risks and tolerability endpoints. Externally methods allowing for quantitative assessments have been developed which allow uncertainty around observed effects to be incorporated however there are still some key limitations. This work presents two methods developed and piloted internally to incorporate not only uncertainty in the observed data but also the correlations between endpoints, the ranges of weights and also target values for each endpoint. This then allows for probabilistic statements about benefit-risk profile to be made and the key components, together with their uncertainty, to be pictorially presented.

PS7E
Statistical Considerations in Evaluating Imaging-Based Devices

Lincoln 5

Organizer(s): *Jincao Wu, FDA/CDRH; Jingjing Ye, FDA; Alicia Toledano; Jeffrey Joseph, Theorem Clinical Research*
Chair(s): *Jincao Wu, FDA/CDRH*

Imaging devices are valuable technologies for primary diagnosis or as an aid to diagnosis for disease screening, diagnosis work-up, monitoring, quantitative biomarker measurement, etc. These imaging devices include radiological techniques to identify abnormalities (e.g., mammography for breast cancer) and digital slides in pathology that allow the pattern recognition and visual search (e.g., tissue slide stained with Her2 for gastric cancer). The evaluation of these devices requires unique analytical and clinical studies. For example, study design and analysis typically needs to address variability in image interpretation by reader. Also, when reading cases in two modalities, the second reading of a case can be affected by memory of the case from the first reading. In this session, statistical considerations on study design and analysis will be discussed among academic, industry and FDA researchers.

Effect of Location Bias in MRMC ROC Studies
Nancy Obuchowski, Cleveland Clinic Foundation; Lucy McGowan, Vanderbilt University; Jennifer Bullen, Cleveland Clinic Foundation

Location bias occurs when a reader detects a false lesion (e.g., non-cancerous lesion) in a subject with disease and the falsely detected lesion is considered a true positive. This can occur in two ways: the reader does not detect the true lesion and only sees false lesions (type I), or the reader detects both the true lesion and false lesions but assigns greater suspicion to the false lesions (type II).

In this study, we examined the effect of location bias in two large MRMC ROC studies, a breast cancer and lung cancer screening study. Our study had five objectives:

1. Examine the frequency of location bias by reader and modality

2. Observe the distribution of confidence scores for each implemented method

3. Identify factors associated with the prevalence of location bias

4. Compare readers' AUC estimates using the different methods

5. Examine the association between the

effect size (difference in readers' AUC between the two modalities) and the difference in the frequency of location bias of the modalities

Location bias occurred in 15–20% of cases overall, but varied among readers from 3–26%. Type I bias was more common in the breast cancer screening study, while type II bias was more common in the lung cancer screening study.

The ROI-ROC and FROC methods are two approaches for correcting for location bias. The methods correct for type I bias by using a confidence score of zero for missed lesions (rather than the confidence score of a false positive); they correct for type II bias by using the confidence score assigned by the reader to the true lesion (rather than to a higher rated false positive). The correction to the confidence score is usually larger for type I bias, shifting confidence scores to zero. However, the overall spread of readers' confidence scores was maintained after the correction.

Readers with higher false positive rates had significantly more location bias, particularly type II bias, with correlations of 0.72 for the breast cancer study and 0.54 for the lung cancer study. Readers' sensitivity was not associated with the frequency of location bias.

The magnitude of the correction to the ROC area estimate is strongly correlated with the frequency of the bias. The correction reduces the magnitude of estimate of the ROC area compared with the uncorrected estimate.

When comparing two modalities' ROC areas, the effect size depends on the difference in the frequency of location bias between the two modalities. When the difference in frequency of bias is small, the effect size is similar whether the location bias is corrected for or not. However, when the frequency of location bias is dissimilar, failure to correct for the location bias favors the modality with higher FPR. In some instances, correcting for the bias increased the effect size, while in others the correction decreased the effect size or reversed the direction of the effect.

Although issues associated with location bias have been described previously and statistical methods to correct for the bias have been available for some time, investigators of MRMC studies seldom discuss the bias or how they handled the bias in their analysis. We conducted a small review of published MRMC studies. We reviewed 27 articles published in Radiology and 34 in AJR in 2014. In 17 of these studies, we felt the potential for location bias existed. One study used FROC, seven used some form of ROI-specific ROC, and nine used neither method. In general, we found that the methodology for addressing the detection of multiple lesions and potential location bias was seldom addressed.

Others (Dendumrongsup 2014) have also addressed the poor quality of data reporting in MRMC studies. They found that the large majority of studies included less than 10 readers. A small number of readers can easily skew study results when location bias is not taken into account, as we presented in our study.

Our study demonstrates the necessity of adjusting for location bias to obtain statistically valid results. From a clinical standpoint, however, the necessity may vary. In the breast cancer example, the current practice dictates that a radiologist correctly locate a potentially cancerous lesion to be biopsied. If the radiologist locates a lesion that is not cancerous, but present in a cancerous breast, it does not help the patient, and in fact can be detrimental. In the lung cancer study, however, if a suspicious lesion is detected on an x-ray, a CT scan of both lungs is performed. In this case, the radiologist's ability to correctly locate the lesion is less important, and perhaps location-bias corrections are less necessary.

In conclusion, location bias can have an impact on the results of a study, particularly if only a few readers are used or if the modality influences the frequency of location bias. In order to avoid spurious results, location bias adjustment is recommended.

Challenges in Digital Pathology and Its Recent Development

Yuying Jin, FDA/CDRH; Meijuan Li, FDA

Advancement in computer technology enables the improved management of image-based information from a digital pathology slide. The digital pathology converts glass slides into digital slides whose pattern can be recognized, searched and analyzed on a computer monitor. There are several challenges in evaluating such device. Reader is often a large source of variability for device of digital pathology. How to assess it appropriately is one of the challenges, in particular models selected for the study that assess reader variability may have an impact on the results of reader variability, as well as other sources of variability included in the study. In addition, the proposed device may have very different technology from the existing device, how to evaluate the performance in a meaningful way is not straightforward. In this presentation, statistical considerations on the study design and analyses will be discussed based on our experience and current thinking.

Statistical Considerations in Reviewing Radiological Imaging Devices at CDRH
Qin Li, FDA/CDRH

Radiological imaging devices such as ultrasound, CT, or mammography is a screening, diagnostic, and monitoring tool in providing visual assistance and guidance to the radiologists/physicians in identifying abnormalities such as breast cancer. Evaluation of the clinical performance of such devices often involves the reader's interpretation. In this talk, a typical study called multi-reader multi-case (MRMC) study is discussed from a regulatory reviewer's perspective. Common issues in study design and analysis encountered when reviewing such devices will be summarized and presented.

PS7F

Use of Phase 2 Interim Analysis to Expedite Drug Development Decisions

Lincoln 6

Organizer(s): *Jenny Huang, Genentech; Qi Xia, Genentech; Qin Li, FDA/CDRH; Norberto Pantoja-Galicia, FDA*
Chair(s): *Jenny Huang, Genentech; Qi Xia, Genentech*

This session aims to discuss the systematic utilization of interim efficacy analyses from comparative phase II trials to expedite the phase III development decisions, a practical approach to shorten the drug development timeline and reduce the development cost, without complicating the trial design and compromising the sponsor's ability to identify gaps in knowledge and thoughtfully design the phase III trial.

We will first examine the theoretical basis as well as the empirical evidence from 35 Roche/Genentech oncology trials for using the phase II interim analysis to facilitate earlier development decisions that is consistent with the final phase II readout, and then go through examples from recent Roche/Genentech oncology trial experience to address issues related to the benefit, cost and implementation details.

Use of Phase 2 Interim Analysis to Expedite Drug Development Decisions
Qi Xia, Genentech

A randomized Ph2 oncology trial with progression-free survival (PFS) as primary endpoint typically takes around two years to gather enough data for a relatively robust Go/No Go decision making to Ph3. In order to expedite the drug development timeline and reduce the development cost, we proposed a consistent decision-making strategy to enable earlier development decisions without complicating the trial design and compromising the sponsor's ability to identify gaps in knowledge and thoughtfully design the phase 3 trial.

We will first examine the theoretical basis and empirical evidence from Roche/Genentech oncology trials for using the phase 2 interim analysis to facilitate earlier development decisions that is consistent with the final phase 2 readout. The predictive probability method will be applied to determine the early decision criteria incorporating different levels of historical knowledge into the prior distribution. The false Go/No Go risks associated with the early decision will be characterized. Last, we will address issues related to the benefit, cost and implementation details when the strategy is being applied to trials in real world from Roche/Genentech experience.

FDA Guidance and Regulatory Experience on Enrichment Design
Yuan Shen, FDA/CDER/OTS/OB; Raji Sridhara, FDA/CDER/OTS/OB

In the era of personalized medicine and targeted therapy, it is important that the drug development process is conducted efficiently. It is therefore essential to adapt during trial based on the interim results observed to stop for futility, efficacy or modify the study by increasing sample size, or focus the remainder of the study on a subpopulation. We will present the FDA guidance on enrichment design and regulatory experience in these type of adaptive designs.

American Statistical Association Conference on

STATISTICAL PRACTICE

Calling all statistical practitioners!

Are you a data analyst, researcher, or data scientist?

CSP 2016 will bring you a full program of concurrent sessions, as well as short courses, tutorials, practical computing demonstrations, exhibits, poster sessions, and the Career Service. Learn new statistical methodologies and best practices in statistical analysis, design, consulting, and programming.

Hone your skills by following a track, or choose from each area:

- Communication, Impact, and Career Development
- Data Modeling and Analysis
- Big Data Prediction and Analytics
- Software, Programming, and Graphics

KEY DATES:

Early Registration Opens: October 1, 2015
Regular Registration Opens: January 6, 2016
Hotel Reservation Deadline: January 19, 2016
Registration Deadline: February 4, 2016

Find out more at *www.amstat.org/csp.*

February 18–20, 2016 – San Diego, CA

CSP