Volume 25, No. 2 • Winter 2018

BIOPHARMACEUTICAL REPORT

Chair: Heather Thomas **Editors:** Amy Xia, Junyuan Wang, Jeff Maca

CONTENTS

Note from the editors

Welcome to the final issue of the Biopharmaceutical (BIOP) Report for 2018! In this issue, we will review some of the productive and exciting accomplishments in 2018 for the Biopharmaceutical Section that will be continued into the following year. This issue's featured article was written by **Ilya Lipkovich** of Eli Lilly and Company and **Alex Dmitrienko** of Mediana Inc, which describes issues and examples with Exploratory and confirmatory subgroup analysis in clinical trials.

This issue also presents updates on some other Biopharmaceutical Section activities, such as a summary from the **Juliet Ndukum** on the Biopharmaceutical Section mentoring program activities from the launch of the program to the present day. There is also an announcement for the upcoming 2019 ASA-Biopharmaceutical section nonclinical biostatistics conference to be held next summer. We also have a few updates from the Section working groups, including updates of the Alzheimer's disease Scientific Working from **Hong Liu-Seifert** and **Steve Wilson.**

The Biopharmaceutical report editors are looking forward to continuing to help provide information for our members in 2019, and we also welcome feedback and suggestions to help improve the BIOP Report. We hope you enjoy reading this issue, and wish everyone all the best in the upcoming and exciting 2019!

EXPLORATORY AND CONFIRMATORY SUBGROUP ANALYSIS IN CLINICAL TRIALS

Ilya Lipkovich, Eli Lilly and Company and Alex Dmitrienko, Mediana Inc

1. INTRODUCTION

Evaluating heterogeneity of treatment effects in clinical trials remains one of the most challenging tasks of the drug development process. On one hand, the goal of a pivotal Phase III clinical trial is to ascertain the treatment effect in the target population of patients in order to ultimately support appropriate drug labeling. Clinical trials are rarely designed with a clear expectation that the beneficial effect of treatment may be limited to a subpopulation. On the other hand, modern precision (a.k.a. personalized) medicine assumes customization of therapies with respect to individual patient characteristics, which means that the "one size suits all" approach is no longer tenable (Ruberg at al., 2010). Identifying treatment effect heterogeneity is especially warranted in situations when (1) the overall treatment effect is driven by a subpopulation and/or (2) an unacceptable safety signal is detected within a certain subpopulation.

Several regulatory guidelines recently released by the U.S. Food and Drug Administration (FDA) and European Medicines Agency (EMA) provide much useful information on the general topic of subgroup evaluation in late-stage clinical trials. General biomarker-related considerations, including biomarker-driven designs, are presented in the FDA guidance on enrichment strategies (FDA, 2012). The EMA guideline on subgroup analysis (EMA, 2014) discusses approaches to investigating subgroup effects in clinical trials with emphasis on an exploratory setting. The confirmatory setting is considered in the FDA and EMA guidelines on multiplicity issues in clinical trials (FDA, 2017; EMA, 2017). The two guidance documents focus on pivotal trials with a few pre-specified patient populations. A similar setting is assumed in the FDA and EMA guidelines on adaptive designs (EMA, 2007; FDA, 2018) when data-driven rules for population selection are described.

Following the framework in a survey of current industry practices (Mayer et al., 2015) we consider the following four types of subgroup analyses in clinical trials:

- Confirmatory.
- Exploratory (in the narrow sense).
- Post-hoc.
- Biomarker and subgroup discovery.

This taxonomy reflects current clinical practice rather than what may be desired in the "ideal world."

Type 1 (Confirmatory subgroup analysis): This includes strategies that are preplanned, i.e., defined prior to data unblinding, and where the number of hypotheses (that may include various subgroups and the overall population) is relatively small. Confirmatory subgroup analysis is limited to biomarkers/subgroups previously identified (perhaps in earlier Phase II studies) and where the Type I error is required to be controlled in the strong sense.

Type 2 (Exploratory subgroup analysis in the narrow sense): This includes strategies specified in the exploratory analysis section of statistical analysis plans. The number of evaluated biomarkers is relatively small, typically limited to known prognostic variables (some of them are included as stratification covariates in the primary analysis). Often it is conducted using multi-stage strategies where the first group of biomarkers is selected by fitting separate regression models with one biomarker at a time. Then each selected continuous biomarker is examined further by choosing an optimal cutoff to define patient subgroups. The Type I error rate is typically controlled only for some elements of the multi-stage strategy, e.g., selection of the biomarker-specific cutoffs.

Type 3 (Post-hoc subgroup analyses): This covers subgroup investigations that are unanticipated prior to

data unblinding and therefore are not pre-specified in a statistical analysis plan. Typically, these are subgroups with unanticipated post-hoc findings after data unblinding that may have trigged regulatory inquiries or subgroups that raise regulatory or sponsor's concerns after drug approval. The Type I error rate is typically only partially controlled within the selected set subgroups (ignoring the fact that the set was chosen in a data-driven manner).

Type 4 (Biomarker and subgroup discovery): This includes data mining of large sets of available candidate biomarkers and is not limited to those anticipated as potential predictors prior to data unblinding. Traditionally, false positive rates are not controlled and external validity is established using cross-validation or replication based on independent data sets. In some cases, however, some form of the overall Type I error rate control or the false discovery rate control is incorporated (this approach will be advocated later in this article).

The main goal of this article is to touch upon key issues arising in subgroup analysis and statistical methods used in this area. Section 2 provides an overview of general principles of exploratory subgroup analysis in the broad sense that covers Types 2, 3 and 4. Type 1 (confirmatory subgroup analysis) will be discussed in Section 3. For an in-depth review of statistical approaches to exploratory and confirmatory subgroup investigations, see for example Ondra et al. (2016), Henderson (2016), Lipkovich, Dmitrienko and D'Agostino (2017) and Lamont et al. (2018).

2. EXPLORATORY SUBGROUP ANALYSIS

The term "exploratory subgroups analysis" is often used to cover a variety of situations when subgroups are evaluated without strictly controlling the Type I error rate. In this section, we will refer to exploratory subgroup investigation in a broad sense, i.e., Types 2, 3 and 4 of subgroup analysis defined in the Introduction.

Key principles of exploratory subgroup analyses

The EMA guideline on subgroup analysis (EMA, 2014) provides useful points to consider when planning subgroup investigation activities. Briefly, the guideline attempts to discourage trial sponsors from making wrong decisions at two extremes. The guideline warns against dismissing subgroup analysis, which is seen in the context of current practices often "creating disincentive to properly plan the investigation of subgroups," and also warns against "reckless" subgroup analysis (that does not exercise caution). In particular, the EMA guideline encourages discussion about potential subgroups at the trial design stage arguing that done properly, "this should minimize the need for data-driven investigations, relying instead on a well-reasoned pre-specified strategy."

However, the guideline does not seem to recognize that a data-driven subgroup investigation strategy can also be principled (and even pre-specified) and the guideline fails to connect this principled approach with a wealth of relevant methods that have been developed in the areas such as statistical learning, causal inference and multiple testing. Indeed, we have witnessed a surge of publications on data-driven subgroup analysis coming from a cross-fertilization of these areas. This is not surprising if we take a view that data-driven subgroup analysis is a special case of model selection in the presence of a large number of biomarkers. In this situation, one of the challenges of modeling is that it aims at studying the causal treatment effect at the individual patient level, which is unobservable, except with a cross-over design.

The key principles of principled subgroup analysis can be extracted from the fields mentioned above. Here we present a brief list. A detailed discussion of these topics can be found in Lipkovich, Dmitrienko and D'Agostino (2017).

- Applying complexity control to prevent data overfitting and selection bias, e.g., bias due to selecting the best patient subgroup from a large set of candidate biomarkers (patient characteristics) and associated cutoffs. Tuning parameters controlling the subgroup search process often need to be determined in a data-driven fashion, e.g., via cross-validation.
- Evaluating the Type I error rate for the entire subgroup search strategy, e.g., by using resampling under the null hypothesis of no subgroup effects. Subgroup analyses are often performed in clinical trials using a multi-stage strategy as described in the Introduction where a multiplicity correction is applied to the last stage but is not applied at earlier stages.

- Obtaining "honest" estimates of the treatment effect within identified subgroups, which are expected if evaluated in an independent (future) data set. In the absence of independent data this can be approximated by using resampling methods or Bayesian model averaging/Empirical Bayes. Again, uncertainty associated with the subgroup identification should be taken into account.

Complex subgroup identification strategies with built-in multiplicity and complexity control have been successfully applied both prospectively and retrospectively across Phase II and III development programs. *Prospective* strategies rely on selecting biomarkers in a Phase II trial and use them to set up enrichment designs in Phase 3 trials. For example, this strategy was employed in the LAVOLTA I and LAVOLTA II trials which will be described in Section 3. Retrospective exploratory subgroup analyses are often conducted in failed Phase III trials to help identify one or more promising subgroups with a beneficial treatment effect. A retrospective subgroup identification strategy was used to discover subgroups with enhanced treatment effect based on 27 baseline covariates in the ATTAIN program for the treatment of nosocomial pneumonia, see Dmitrienko et al. (2015).

Typology of exploratory subgroup analysis methods

Following Lipkovich, Dmitrienko and D'Agostino (2017), we briefly outline four classes of methods that emerged in the recent literature on data-driven subgroup analysis.

Global outcome modeling

This approach assumes that outcome models in the treatment and control arms are estimated from the clinical trial data. A single regression model incorporating both main (prognostic) effects and treatment by covariate interactions (predictive effects) or separate models by treatment arm may be used. Constructing patient subgroups typically requires multi-stage procedures. As an example, the global outcome model may be estimated at the first stage using a "black box" model (Neural network, Random forest, Gradient boosting) and used at the second stage to compute hypothetical individual treatment differences. These are, in turn, used as outcomes and modeled using appropriate predictive modeling methods, e.g., classification and regression trees, to select biomarkers predictive of treatment response and associated patient subgroups. The Virtual twin method introduced in Foster et al. (2011) serves as an example of global outcome modeling strategies.

Global treatment effect modeling

Approaches of this class obviate the need to fit prognostic effects that "cancel out" in the course of modeling. As a result, modeling may be more robust as it is not prone to misspecification of prognostic effects. Some methods in this class rely on the machinery of classification and regression trees while replacing the usual splitting criteria which encourage splits resulting in the largest reduction in node impurity with splits aimed at maximizing the treatment by split interaction. An example of a global treatment effect modeling strategy is the Interaction tree method introduced in Su et al. (2009).

Modeling individual treatment regimes (ITR)

Broadly, this class includes any approach that determines, based on patient level data, a rule determining for any given patient which candidate treatment is more likely to result in a better response as compared to other treatments. An important subclass of ITR methods casts the task of identifying optimal treatment rule as a classification problem where the goal is predicting the sign of the hypothetical treatment difference (i.e., if positive, assign a patient to the experimental treatment and, if negative, to the control). Methods of this class were pioneered by the Outcome-Weighted Learning approach of Zhao et al. (2012).

Local modeling (direct subgroup search)

The last class of subgroup evaluation methods focuses on a direct search for treatment-by-covariate interactions and then selecting subgroups with desirable characteristics, e.g., subgroups with an improved treatment effect. This approach obviates the need to estimate the response function over the entire covariate space and focuses on identifying specific regions with a large differential treatment effect. Unlike brute-force search methods used in computer science, local modeling methods often incorporate complexity control and multiplicity adjustments. Examples include extensions of "bump hunting" to subgroup analysis by Chen et al. (2015) and SIDES and related methods (see Lipkovich et al., 2011; Lipkovich and Dmitrienko, 2014; and Lipkovich et al., 2017).

3. CONFIRMATORY SUBGROUP ANALYSIS

As explained in the introduction, confirmatory subgroup analysis methods focus on problems with a small set of pre-specified subgroups in pivotal clinical trials, which includes the overall trial population and several prospectively defined subpopulations. In this section, we will describe the general class of multi-population trials and a number of relevant topics such as the choice of multiplicity adjustments in traditional multi-population trials and decision rules in adaptive trials with population selection.

Multi-population trials

Confirmatory subgroup analysis commonly arises in pivotal trials aimed at the development of targeted therapies. In this case, one or more subsets of the overall population are defined using binary classifiers derived from baseline patient characteristics (biomarkers) and where the treatment benefit is expected to be stronger in these subpopulations than in the overall population. The subpopulations are often referred to as target subpopulations.

The APEX trial (Cohen et al., 2016) serves an example of a multi-population trial with two target subpopulations. This trial was conducted in the population of patients at risk for venous thrombosis and was designed to evaluate the efficacy and safety of betrixaban compared to an active control. The target subpopulations were defined using the baseline values of two variables/biomarkers (patient's age and D-dimer level) which were believed to be predictive of treatment response. Incorporating the subpopulations into the primary analysis would help the trial sponsor better characterize the efficacy profile of betrixaban.

Multiplicity adjustments

A key feature of any multi-population trial is that an effectiveness claim can be made in the overall trial population as well as in any of the target subpopulations. This leads to Type I error rate inflation and thus to control the overall Type I error rate with respect to the corresponding null hypotheses of no effect, a multiplicity adjustment must be applied.

A comprehensive overview of multiplicity adjustment strategies utilized in pivotal clinical trials is provided in Dmitrienko and D'Agostino (2013, 2018). Numerous multiple testing procedures are available (e.g., nonparametric, semiparametric or fully parametric procedures) to protect the overall Type I error rate in multi-population trials and it is important to select the most appropriate procedure for a given trial. The process of carefully evaluating the available multiplicity adjustment options can be facilitated by an application of the Clinical Scenario Evaluation (CSE) approach (Benda et al., 2010).

To illustrate the importance of a thorough evaluation of multiplicity adjustments, we will consider the APEX trial with three pre-defined patient populations. A simple fixed-sequence testing approach was pre-planned in this trial and the patient populations were to be tested sequentially beginning with the smallest subpopulation. The treatment effect was not significant within this subpopulation and due to the inflexible testing approach, the trial's overall outcome was declared negative. This testing strategy serves as an example of a multiplicity adjustment that imposes unnecessary restrictions on the population-specific significance tests. Patient populations ought to be treated as interchangeable rather than hierarchically ordered. For this reason, multiplicity adjustments that rely on a flexible data-driven testing sequence (e.g., the Hochberg procedure) are recommended. Secondly, instead of focusing on a single multiplicity adjustment, the CSE approach encourages trial sponsors to select a set of applicable candidate adjustments and carefully evaluate them to identify the most efficient adjustment that performs well under a broad range of treatment effect assumptions (including the scenarios where the biomarkers of interest are non-informative, i.e., they do not predict treatment benefit). Examples of CSE-based evaluations of multiplicity adjustments in multi-population trials can be found in Dmitrienko and Paux (2017).

Decision-making framework

While multiplicity adjustments provide a foundation for valid statistical inferences in multi-population trials, they do not guarantee a logically consistent set of conclusions. An additional set of conditions should be imposed to facilitate the interpretation of subgroup analysis results in trials with several patient populations.

This point can be illustrated using the SATURN trial that investigated the use of erlotinib as maintenance therapy in patients with advanced non-small-cell lung

cancer (Cappuzzo et al., 2010). The primary analysis in this trial was performed in the overall population as well as a subpopulation of patients with a positive EGFR (epidermal growth factor receptor) immunohistochemistry status. Progression –free survival was the endpoint evaluated in the primary analysis in both patient populations. The treatment effect turned out to be significant in the overall population as well as in the subset of EGFR-positive patients. However, as pointed out by Rothmann et al. (2012), it is premature to conclude that erlotinib provided benefit in the overall population since it is possible that the significant outcome in the overall population was driven by a strong treatment effect within the target sub population.

To ensure appropriate inference in multi-population settings, Millen et al. (2012) introduced the influence and interaction conditions. The influence condition states that in order to claim effectiveness in the overall population, a beneficial treatment effect in this population must not be limited to the target subpopulation. This implies that the treatment benefit in the overall population is not explained solely in terms of a strong treatment effect in the subpopulation. Furthermore, the interaction condition applied to settings where the trial sponsor would like to make effectiveness claims simultaneously in the overall population and target subpopulation. If the influence condition is satisfied, the interaction condition states that in order to support both effectiveness claims, the treatment effect in the target subpopulation should be appreciably greater than the treatment effect in the complement of the subpopulation.

The original influence and interaction conditions were formulated using a simple frequentist approach and were later expanded to incorporate Bayesian arguments. The decision-making framework based on these conditions has been applied to several multi-population trials with traditional and adaptive designs.

Adaptive population selection trials

So far we have focused on multi-population trials with a fixed design where the total sample size or target number of events in a trial is pre-specified. The multi-population framework has been successfully extended to designs with data-driven decision rules. These are known as either adaptive population selection designs or adaptive enrichment designs (FDA, 2018). In what follows, we will discuss well-established adaptive designs aimed at evaluating treatment benefits in a set of subpopulations that are defined at the trial design stage as well as more advanced designs that are built around data-driven subpopulations selected at an interim analysis.

Beginning with adaptive designs that rely on a set of pre-specified subpopulations, a multi-population adaptive trial with a single interim analysis can be designed to evaluate the treatment effect within each subpopulation at the interim look and identify the most promising sub-populations. The final analysis is then conducted within the selected sub-populations using the data collected before and after the interim analysis. In addition, the number of patients or events can be appropriately increased in the overall population or within a target sub-population if the treatment effect at the final analysis is projected to be borderline non-significant. Most commonly, trials with either adaptive subpopulation selection or other data-dependent rules rely on the combination function principle, which guarantees overall Type I error rate control. A detailed example of an adaptive multi-population trial with flexible decision rules is provided in Brannath et al. (2009).

For evaluating the overall population, the resulting adaptive approach offers several advantages compared to a traditional design. As stated in the FDA guidance on adaptive designs (FDA, 2018), adaptive designs are likely to provide a power advantage over traditional designs In addition, the adaptive approach offers a certain level of protection against incorrect selection of target subpopulations or misspecification of the treatment effect assumptions in Phase III development programs. It is well known that it is challenging to reliably assess the predictive properties of a promising biomarker in Phase II trials, mostly due to their small size. For example, a continuous biomarker (serum periostin) was developed as a predictor of treatment response in asthma populations and showed considerable promise in a Phase II trial in patients with uncontrolled asthma (Corren et al., 2011). This Phase II trial was conducted to investigate the efficacy of lebrikizumab and a strong treatment effect was detected in the periostin-high subgroup (biomarker-positive patients) whereas there was no evidence of efficacy in the complementary subgroup (biomarker- negative patients). Two large Phase III trials using traditional designs with a fixed sample size (LAVOLTA I and LAVOLTA II) were conducted to confirm the subgroup findings (Hanania

et al., 2016). However, the enhanced efficacy signal in the biomarker-positive subgroup was not observed consistently. In particular, a significant treatment effect was found in the biomarker-positive subgroup for the LAVOLTA I trial but the treatment difference in the subgroup was borderline non-significant in the other trial.

A natural extension of adaptive designs with fully pre-specified subpopulations is a class of more advanced designs where patient subpopulations are partially defined at the beginning of the trial (i.e., predictive biomarkers are known but the rules for defining patient subgroups may not be specified). Using the lebrikizumab example with a continuous biomarker, consider the problem of selecting the subset of biomarker-positive patients for pivotal phase III trials. Cut points for continuous biomarkers are typically estimated from small Phase II trials. Since Phase III trials tend to be larger than Phase II trials, an alternative approach would be to introduce an interim analysis in a Phase III trial and use the interim analysis results to identify an optimal cut point. The subpopulation corresponding to this cut point could then be analyzed at the final look as if it was pre-planned at the trial design stage.

One of the first attempts to build designs with data-driven subpopulations for pivotal trials was proposed by Freidlin and Simon (2005); however, the proposed approach discarded the data collected before the interim look. More recently, more powerful methods have been developed to perform valid inferences in data-driven subgroups by pooling the subgroup data before and after the interim look (see Graf et al., 2019). This general inferential framework can be extended to more complex settings. For example, the trial sponsor can identify a set of candidate biomarkers and apply a principled subgroup identification method (see Section 2) to choose the strongest predictor of treatment benefit and associated target subpopulation at the interim analysis. The treatment effect will then be estimated within the overall trial population as well as in the selected subpopulation using an appropriate multiplicity adjustment.

4. DISCUSSION

This article provides an overview of key considerations and methods in the evaluation of subgroup effects in late-stage clinical trials. Multiple patient subgroups are typically examined in Phase II and Phase III clinical trials and it is critical to identify appropriate statistical methods that are aligned with the goals of subgroup analyses, e.g., data-driven or confirmatory subgroup analyses.

In the context of data-driven subgroup analysis, depending on the scope of subgroup investigation, we distinguish among the traditional exploratory analysis in clinical trials, post-hoc analysis and subgroups/biomarker discovery. Multiple statistical methods have been proposed recently for unplanned subgroup analysis and we would like to emphasize the importance of a principled approach to subgroup exploration. This class of methods is the result of a cross-fertilization of efforts form machine learning, causal inference and multiple testing. Within the principled subgroup analysis framework, complexity and error rate control should be implemented using resampling methods to allow the trial's sponsor to account for the uncertainty in complex multi-stage biomarker/subgroup search strategies. Treatment effects within the identified subgroups (or the improvement in overall outcomes resulting from applying estimated individual treatment regimes) should be estimated by methods that account for selection bias.

Multiplicity adjustments play a key role in confirmatory subgroup analysis settings (in fact, multiplicity adjustments are mandatory in pivotal Phase III trials with multiple patient populations). When comparing the available multiplicity adjustment options in a multi-population trial, it is important to identify an adjustment that performs best in the context of a given trial and is also robust against deviations from the original treatment effect assumptions. In addition, increasingly more sophisticated adaptive designs are available to support the investigation of subgroup effects in pivotal trials. These designs incorporate flexible data-driven decision rules, such as discontinuation of patients or selection of the best subpopulations from a set of candidate subpopulations, and enable valid statistical inferences that do not compromise Type I error rate control.

The use of appropriate designs and methods will allow for a better understanding of treatment response, for more efficient regulatory review and will ultimately provide greater benefit to the patients through reliable knowledge in tailored therapeutics.

ACKNOWLEDGEMENT

The authors thank Anthony Zagar for reviewing the article and making valuable comments.

REFERENCES

1. Benda N, Branson M, Maurer W, Friede T. Aspects of modernizing drug development using clinical scenario planning and evaluation. *Drug Information Journal*. 2010; 44:299-315.

2. Brannath W, Zuber E, Branson M, Bretz F, Gallo P, Posch M, Racine-Poon A. Confirmatory adaptive designs with Bayesian decision tools for a targeted therapy on oncology. *Statistics in Medicine*. 2009; 28:1445-1463.

3. Cappuzzo F et al. Erlotinib as maintenance treatment in advanced non-small-cell lung cancer: A multicentre, randomised, placebo-controlled Phase 3 study. *Lancet Oncology*. 2010; 11:521-529.

4. Chen G, Zhong H, Belousov A, Viswanath D. PRIM approach to predictive-signature development for patient stratification. *Statistics in Medicine*. 2015; 34:317-342.

5. Cohen AT et al. Extended thromboprophylaxis with betrixaban in acutely ill medical patients. *New England Journal of Medicine*. 2016; 375:534-544.

6. Corren J et al. Lebrikizumab treatment in adults with asthma. New England Journal of Medicine. 2011; 365:1088-1098.

7. Dmitrienko A, D'Agostino RB. Tutorial in Biostatistics: Traditional multiplicity adjustment methods in clinical trials. *Statistics in Medicine*. 2013; 32:5172-5218.

8. Dmitrienko A, Lipkovich I, Hopkins A, Li YP, Wang W. Biomarker evaluation and subgroup identification in a pneumonia development program using SIDES. In *Applied Statistics in Biomedicine and Clinical Trials Design*, Chen Z, Liu A, Qu Y, Tang L, Ting N, Tsong Y (editors). Springer, New York, 2015; 427-466.

9. Dmitrienko A, Paux G. Subgroup analysis in clinical trials. *Clinical Trial Optimization Using R*. Dmitrienko A, Pulkstenis E (editors). Chapman and Hall/CRC Press, New York, 2017.

10. Dmitrienko A, D'Agostino RB. Multiplicity considerations in clinical trials. *New England Journal of Medicine*. 2018; 378:2115-2122.

11. EMA (European Medicines Agency). Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. 2007.

12. EMA (European Medicines Agency). Guideline on the investigation of subgroups in confirmatory clinical trials. 2014.

13. EMA (European Medicines Agency). Guideline on multiplicity issues in clinical trials. 2017.

14. FDA (U.S. Food and Drug Administration). Guidance for industry: Enrichment strategies for clinical trials to support approval of human drugs and biological products. 2012.

15. FDA (U.S. Food and Drug Administration). Guidance for Industry: Multiple endpoints in clinical trials. 2017.

16. FDA (U.S. Food and Drug Administration). Guidance for Industry: Adaptive design clinical trials for drugs and biologics. 2018.

17. Freidlin B, Simon R. Adaptive signature design: An adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients. *Clinical Cancer Research*. 2005; 21:7872-7878.

18. Foster JC, Taylor JMC, Ruberg SJ. Subgroup identification from randomized clinical trial data. *Statistics in Medicine*. 2011; 30:2867-2880.

19. Graf AC, Wassmer G, Friede T, Gera RG, Posch M. Robustness of testing procedures for confirmatory subpopulation analyses based on a continuous biomarker. *Statistical Methods in Medical Research*. 2019. To appear.

20. Hanania NA et al. Efficacy and safety of lebrikizumab in patients with uncontrolled asthma (LAVOLTA I and LAVOLTA II): Replicate, Phase III, randomised,

double-blind, placebo-controlled trials. *Lancet Respiratory Medicine*. 2016; 4:781-796.

21. Henderson NC, Louis TA, Wang C, Varadhan R. Bayesian analysis of heterogeneous treatment effects for patient-centered outcomes research. *Health Services and Outcomes Research Methodology.* 2016; 16:213-233.

22. Lamont A, Lyons MD, Jaki T, Stuart E, Feaster DJ, Tharmaratnam K, Oberski D, Ishwaran H, Wilson DK, Horn MLW. Identification of predicted individual treatment effects in randomized clinical trials. *Statistical Methods in Medical Research.* 2018; 27:142-157.

23. Lipkovich I, Dmitrienko A, Denne J, Enas G. Subgroup identification based on differential effect search (SIDES): A recursive partitioning method for establishing response to treatment in patient subpopulations. *Statistics in Medicine.* 2011; 30:2601-2621.

24. Lipkovich I, Dmitrienko A. Strategies for identifying predictive biomarkers and subgroups with enhanced treatment effect in clinical trials using SIDES. *Journal of Biopharmaceutical Statistics.* 2014; 24:130-153.

25. Lipkovich I, Dmitrienko A, Patra K, Ratitch B, Pulkstenis E. Subgroup identification in clinical trials by Stochastic SIDEScreen methods. S*tatistics in Biopharmaceutical Research.* 2017; 9:368-378.

26. Lipkovich I, Dmitrienko A, D'Agostino RB. Tutorial in Biostatistics: Data-driven subgroup identification and analysis in clinical trials. *Statistics in Medicine.* 2017; 36:136-196.

27. Mayer C, Lipkovich I, Dmitrienko A. Survey results on industry practices and challenges in subgroup analysis in clinical trials. S*tatistics in Biopharmaceutical Research.* 2015; 7:272-282.

28. Millen B, Dmitrienko A, Ruberg S, Shen L. A statistical framework for decision making in confirmatory multi-population tailoring clinical trials. *Drug Information Journal.* 2012; 46:647-656.

29. Ondra T, Dmitrienko A, Friede T, Graf A, Miller F, Stallard N, Posch M. Methods for identification and confirmation of targeted subgroups in clinical trials: A systematic review. J*ournal of Biopharmaceutical Statistics*. 2016; 26:99-119.

30. Rothmann MD, Zhang JJ, Lu L, Fleming TR. Testing in a pre-specified subgroup and the intent-to-treat population. *Drug Information Journal.* 2012; 46:175-179.

31. Ruberg SJ, Chen L, Wang Y. The mean does not mean as much anymore: finding sub-groups for tailored therapeutics. *Clinical Trials.* 2010; 7:574-583.

32. Su X, Tsai CL, Wang H, Nickerson DM, Li B. Subgroup analysis via recursive partitioning. *Journal of Machine Learning Research.* 2009; 10:141-158.

33. Zhao Y, Zheng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association.* 2012; 107:1106-1118. ■

OVERVIEW OF BIOPHARMACEUTICAL SECTION **MENTORING ACTIVITIES FROM LAUNCH TO PRESENT DAY**

BRIEF INTRODUCTION & BACKGROUND

Juliet Ndukum

Mentorship within the American Statistical Association (ASA) began not long ago with the Committee on Applied Statisticians (CAS) being advanced in providing resources for this program to its members. Following the Committee on Applied Statisticians blueprint, the Biopharmaceutical (BIOP) Section created the BIOP Section Mentoring Committee to coordinate mentorship of junior colleagues by experienced statisticians within the section. The mentoring program was introduced by the ASA as added benefit to its members. BIOP Section was the first Section to start the mentoring program following the blueprint created by the committee on applied statisticians (CAS).

The BIOP Section Mentoring Program was launched in 2014. Since its launch, the program has seen participation from over a hundred mentees. For these years the interest has been growing among Biopharm members and potential mentees/mentors. On the other hand, surveys administered by the mentoring committee during each year's program have revealed that some mentoring partners did not have sustained contacts after the initial meeting through the end of the program. In order to have continuity of the program, it is necessary to provide incentives, to encourage volunteers and promote active participations in future years. The program has gained popularity as we have participants from outside USA – mentee/mentor from outside USA. As the years go by, more and more people have participated in the program.

THE FIRST INITIAL YEARS 2014-2016

Amajor Kaur together with Jennifer Gauvin served as chair of CAS for multiple years and were part of the mentoring initiative/pilot. The activities of the first three years (2014-2016) set the stage for the program. These activities included creation of an email account for communication about planned activities to target audience.

Putting together useful information in form of a welcome package to be shared to all mentee/mentor pairs. The document (welcome package) modified from CAS to suit BIOP Section community provides information to mentee/mentor pairs about program expectations. In order to gauge interaction and interest in the program, a survey was created. This survey is administered twice during the program. Another survey also aids in pairing mentees to mentors. A series of meetings and events were planned, and held with the purpose of bringing about awareness of the program to the BIOP Section community. There was a focus on mentoring at JSM 2016 with talks on mentoring. Because of this interest, the committee published an article to highlight the importance of the program to the BIOP Section community in the September edition of *AMSTATNEWS*.

ACTIVITIES OF THE COMMITTEE

The mentoring committee advertises solicits interest in the program by advertisement through various channels e.g., BIOP email/newsletter/business meeting at JSM, JSM mixer, *AMSTATNEWS*, and STAT*tr@k*. Individuals who indicate an interest have to do confirm their BIOPiopharm membership as a first step to participating in the program. Each mentee is matched with a

mentor. The matching process is based on their goals, learning objectives from the mentoring program, working background, and time geographic location of both mentee and mentor. The mentor/mentee pairing process is facilitated by an uptake survey where each interested participant fills out to provides the necessary information on the questions asked in the survey. The committee then uses the feedback from the survey as a guide to provide the best possible matching pairs. The timing of the pairing is such that, both mentor and mentee can have a first meeting during JSM of that year. Once a mentee is matched to a mentor, the new mentor/mentee pair are informed by email with a welcome package provided as attachment. The committee encourages at least one face-to-face meeting for example during JSM or another conference, if possible, that is not withstanding other forms of communication (email, text, talk, teleconferencing) or a combination thereof. During the midyear, a follow-up email is sent to mentor-mentees to keep them engaged in the program and also to remind them about their ongoing communications responsibilities, etc. The committee envisages planning a session at the 2019 FDA Statistics Workshop and improve upon the BIOP website about the program and its activities.

ACTIVITY, EVALUATION & DISSEMINATION

The BIOP mentorship program runs from August to July in the following year where the mentoring committee members matches interested mentees to mentors and continues following up with the matching pairs through the year. After the initial match, participants receive two surveys In March/April following the initial match, while they are in the program, another survey is emailed to participants. The survey is intended to capture progress of the mentoring program, interaction as well as interest. Additionally, the midyear survey administered is intended to evaluate the activities of the program based on the responses of the mentee/mentor pairs. The feedback from the survey is used to identify areas of improvement within the program. Although not a requirement, a mentee/mentor pair may opt to continue with the communication after the mentoring program has ended. Furthermore, a mentee may have interest to be matched with a mentor for the next year of the program. Towards the end of each mentoring session, a final survey is again administered to the participants. As before, it is intended to capture the same attributes, as the previous and identify lapses, if, any, in the program. The information from both surveys are used to identify areas of improvement and innovation within the program. These lapses, again, provide room for improvement and innovation. At the end of the program each year, a report is presented at the Executive Committee of the BIOP Section, a summary published in the BIOP newsletter.

ACTIVITIES POST LAUNCH OF THE PROGRAM

The activities post the initial three years have been varied . The BIOP Section Mentoring Committee is writing this report to create awareness of the program and milestones including successes of the program to BIOP Section and ASA as a whole and our noble profession. The mentor participation in the BIOP Section mentoring program has been low compared to the number of mentees seeking mentorship. To facilitate this high demand for mentorship in our Section, in 2015 and 2016, the BIOP Section Mentoring Committee members, with the consent of mentors, assigned multiple mentees to some individual mentors. The presence of mentors is needed to guarantee continuity of the program. (Mentor participation was low, this year as well.) This year, at JSM 2018, the Executive Committee (EC) brought the idea to the attendees at the BIOP Section Business Meeting and Mixer in Vancouver.

RECOGNITION OF ALL MENTORS TO DATE

By mentoring a junior colleague, the mentor understands the junior colleague's professional development needs, and serves as a valuable resource for the mentee for technical issues and career-related challenges to name a few. A mentor may spend a substantial amount of time to address requests from mentees. Because of the efforts mentors put in and the advice they give to the mentee. the BIOP Section Mentoring Committee would like to thank all senior colleagues who have served on the program as mentors at some point in time since the launch of this program. ■

ALZHEIMER'S DISEASE SCIENTIFIC WORKING GROUP **UPDATES**

Hong Liu-Seifert and Steve Wilson, The Alzheimer's Disease Scientific Working Group

The Alzheimer's Disease Scientific Working group has been making steady progress. The Disease Modification and the Pre-symptomatic Work Stream are currently developing review white papers to help better understand the approaches that have been used/proposed to demonstrate disease modifying effects and to utilize appropriate measures for pre-symptomatic trials.

Due to the urgency of the estimands topic, our Estimand Working Stream has been especially busy. Listed below are some of the key activities of this group:

1. Organizing a symposium at Clinical Trials for Alzheimer's Disease (CTAD) to educate the AD field on the framework of estimand. Key clinical opinion leaders and the FDA Director of Neurology Division participated in the panel discussion;

2. Presenting at the Statisticians in the Pharmaceutical Industry (PSI) and the Alzheimer's Association International Conference (AAIC) on simulation results focusing on the performance of estimands and estimating methods;

3. Organizing a statistical workshop at the AAIC to further educate the broader research community on estimands and specific considerations dealing with intercurrent events: and

4. Planning to host one of the webinars in the DIA estimand series "Getting the questions right," to share learning and best practice associated with cross-disciplinary groups in order to effectively implement and utilize estimands in CTs and in drug development.

Earlier this year we also participated in a session at DIA/FDA Statistics Forum to discuss the challenges and opportunities in AD drug development and research, and what we, as statisticians, can do to "lead the charge." It was a productive discussion between statisticians and clinicians, including regulators, and it clearly demonstrated the commitment coming from cross-sector efforts to fight this devastating disease.

Additionally, working with Richard Zink, we described Alzheimer's Disease research challenges and the activities of the AD_SWG on the latest Biopharmaceutical Sections podcast. ■

PODCAST | Episode 59: Alzheimer's Disease Scientific Working Group

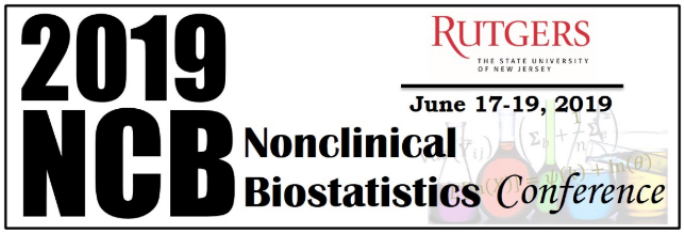
In the latest Biopharmaceutical Section Podcast Hong Liu-Seifert and Steve Wilson discuss the challenges of developing new treatments for Alzheimer's Disease and the formation of the scientific working group.

https://community.amstat.org/biop/home

NONCLINICAL BIOSTATISTICS CONFERENCE 2019

NONCLINICAL BIOSTATISTICS – ADVANCING DRUG DEVELOPMENT FROM DISCOVERY TO COMMERCIALIZATION

The sixth ASA-BIOP Nonclinical Biostatistics Conference is set to take place at Rutgers University in New Brunswick, NJ from June 17 – 19, 2019. The conference is dedicated entirely to nonclinical biostatistics topics with four organized sections for CMC/Manufacturing, drug discovery/biomarkers, safety/pharmacology, and statistical computing and visualizations. Attendees will have ample opportunity to network, share experiences, and discuss current scientific issues with colleagues and leaders in the field. The conference features the ASA Presidential speech from ASA president-elect **Dr. Karen Kafadar** (UVA) and a keynote address from **Dr. Jose Pinheiro** (J&J). In addition, two short courses are offered for the conference : (1) An R shiny tutorial with nonclinical applications with instructors **Max Kuhn** and **Phil Bowsher** (RStudio) and (2) Getting it right: Composition analysis of biological measurements with instructors **Anthony Lonardo** (Lonardo StatReg Associates) and **Juan Jose Egozcue** and **Maribel Ortego** (Dept. Civil and Environment Engineering, Universitat Politecnica de Catalunya). Special programming and events are underway for graduate students, including a student poster contest. The NCB student outreach is currently coordinating with several universities to interact with and educate graduate students on the role of nonclinical biostatistics in the pharmaceutical industry. Further, the 2019 Best Nonclinical Paper Biostatistics Award will be bestowed upon the winner at the conference. Submissions for best paper are accepted up until March, 2019. The organizing committee is preparing a full agenda of invited and contributing speakers that will be selected in the coming months. Look for registration to open in early 2019. On behalf of the conference organizing committee, we look forward to seeing you there! -- Steven Novick



Visit the conference website: ***http://community.amstat.org/biop/events/ncb/index***

Visit the NCB student outreach website: ***http://community.amstat.org/biop/working-groups/ncbwg/students***

Visit the Best non-clinical biostatistics paper website: ***http://community.amstat.org/biop/working-groups/ncbwg/awards***