2021 NCB has officially started! We have two exciting short courses today. The recordings of both courses are available to replay on the conference platform.

* AM short course: Bayesian Regression Trees

Jason Roy, Department of Biostatistics and Epidemiology, Rutgers School of Public Health

Jason Roy of Rutgers presented a short course on Bayesian Regression Trees (BART), comprehensively expounding BART with necessary technical details and hands on examples to enable its use in practice. The main highlights of the course are summarized below.

First a traditional Bayesian analysis and Markov Chain Monte Carlo (MCMC) to carry out inference were introduced. Then a transition from the usual parametric setting to the nonparametric Bayesian methods was explained. As a next step, Bayesian decision trees were elucidated. This included construction of the tree priors via branching processes and the posterior computation through MCMC.

In the second part of the shortcourse, ensembles (collections) of the decision trees were introduced as a main idea underlying BART. Next a step-by-step walkthrough through BART ensemble construction, its naturally ensuing smoothing and shrinkage properties and the Bayesian backfitting was carried out. An example with the R code was introduced to provide guidance on how to use BART in practice. Utility of BART was further demonstrated in two examples from causal inference, which included propensity scores estimation and a treatment effect estimation from observational data in an antiretroviral therapy application.

* PM short course: Bayesian Survival and Joint Models using Rstanarm

Jacqueline Buros Novik, Head of Data and Analytics at Generable Inc.

Jacqueline Buros of Generable presented a short course on fitting Bayesian survival models in R, using recent additions to the rstanarm package. These tools provide for statistical inference in the presence of censoring. The course concentrated on the integration of biomarkers and other longitudinal data sets into a comprehensive statistical platform. Specific recommendations for work flows were discussed, and complete examples of code were included. Emerging code from Github was demonstrated, including the function stan_jm, allowing for joint modeling. These tools are not yet integrated into CRAN, and explicit instruction on their installation was given. This suite of tools was developed and is supported by Generable, and organization supporting the use of Bayesian techniques in clinical trials. Course material is available at https://github.com/generable/ncb2021/.

The second day of the conference started with a welcome by John Kolassa. This brief welcome was followed by a keynote address from Wendy Martinez of the Bureau of Labor Statistics and 2020 American Statistical Association president. The presentation discussed ethical considerations arising from the use of statistical and data science techniques to make decisions that have human impacts. Ethical principles of various statistical organizations and rules produced by various governments were reviewed.

Lori Pfahler discussed the frontier of statistics and data science in CMC. The historical developments of statistics and data science were reviewed, and the integration of these functions into various industrial contexts, and including CMC in the pharmaceutical industry. Considerations of adding value to the organization were reviewed. Interpretability of models was considered. Lori Pfahler recommends a collaborative rather than antagonistic relationship to data science.

Steve Novick discussed replicability and reproducibility in drug discovery. Steve Novick pointed out an alarming pattern of inconsistency in the results of preclinical studies, impacting replicability. He noted that study by treatment interactions can lead to such inconsistency; magnitudes of these interactions may be derived from historical data, using linear mixed-effect models, and these magnitudes may be used to change significance levels. Reproducibility was also discussed.

Guoqing Diao discussed correction for multiplicity for correlated adverse events. Proposed methods are more powerful than Bonferroni correction for signal detection, even in rare events. The proposed method for recovering p values relies on Monte Carlo. The approach is applied to logistic regression with two treatment arms, but it may be extended beyond this context. Utility of the approach was demonstrated in simulation and examples from vaccine and antidepression clinical trials.

Di Cook connected with us asynchronously to discuss data plots. She discussed data plots as statistics, using a grammar of graphics, and that a plot implicitly incorporates a null hypothesis, in that the plot structure defines what comparison would be uninteresting. Identifiable patterns correspond to a rejected null hypothesis. These ideas were applied to dot and bar pots. The tools were presented for various real data sets. Future potential applications include high throughput screening.

Two parallel roundtable discussions were held, on students and early career statisticians, and on industry and FDA interactions. These discussions were wide-ranging, with lively audience discussions.

Round table discussions were followed by two parallel poster sessions. A range of speakers briefly presented their work, with lively discussion.

Misbah Ahmed discussed extension of quantitative decision making tools to CMC, with the aim of estimating probability of success for a particular study. These calculations include specifying a prior distribution, and calculating a marginal probability of success. These tools are applied to transfer of technology between Research and Development to Manufacturing. The presentation results of previously-run software results

Gregory Hather discussed novel assays in preclinical drug development, and addressed the question of minimal assay quality required for the assay to be useful, in order to balance delay of the development

program and failing to cancel a compound early. Gregory Hather presented a model for determining assay cost, and noted that even a not particularly accurate assay can be useful.

Tony Pourmohamad discussed Bayesian tools to reduction of the use of preclinical animal studies, as a tool for avoiding unnecessary animal suffering. A sequential testing approach using Bayes' factor was used. Operating characteristics of the sequential test were discussed, as were operational barriers to using these techniques.

Ya-Ching Matilda Hsieh discussed dose-response curve fitting as a tool for non-clinical modeling, in order to avoid inefficiencies inherent in one-number summaries like the area under the curve. Particular challenges were delays in response, hook effects, and outliers. She proposed a growth curve model with random effects to address the delayed response. The hook effect was alleviated by adding the reflection point to the model. Improved curve fitting was illustrated in examples of several preclinical data sets.

Tim Schofield discussed analytical limits in quality determination in biopharmaceutical manufacturing. Tim discussed a variety of open questions involving fundamental principles of limits, and methods for determining tolerances. Specification limits, release limits, and control limits were discussed, as were impacts of current practices on development.

Recordings for all sessions are available to replay on the conference platform. Look forward to seeing you tomorrow.

Wednesday, Day 3 of the conference, kicked off with a welcome from Don Bennett, who reviewed background of the conference.

Nassim Taleb gave our second keynote address, on the Statistical Consequences of Fat Tails. Prof. Taleb discussed the consequences of various tail behaviors for distributions and noted that distributions with problematic tail behaviors occur in a variety of disciplines, and imply consequences such as nonelliptical multivariate distributions and a lack of relevance of forecasts, and preferred measures of variation. Various other statistical estimation tasks, including regression fitting and principle component analysis, must also be performed differently. He discussed difficulties in identifying the fat-tailed nature of distributions. Prof. Taleb's new book of the same title is freely available at https://researchers.one/articles/20.01.00018

Yiming Peng discussed data analysis techniques applied to cell growth data; he described difficulties in management, organization, and modeling of these data, and discussed roles of various team members on such a project. These considerations were exhibited in the context of a particular biologics manufacturing project. Lessons learned were reviewed.

Yuting Xu discussed machine learning applications to protein engineering. Models are needed in order to predict protein function. Efficiencies here are important because of the time and expense of the needed experiments. Considerations necessary to apply machine learning tools were discussed. Performances of

various machine learning tools were compared, with convolutional neural network outperforming the other methods applied.

Melvin Munsaka discussed spontaneous reporting systems for adverse events. Advantages and disadvantages of reporting systems and existing data sources were discussed. Many of these data sources are publicly available. Tools and algorithms for exploring these data were presented.

Matthew Kay discussed the difficulties in communicating uncertainty, in words and primarily in pictures. Difficulties include systematic biases that people exhibit when interpreting geometric attributes like areas. Recommendations are supported by empirical studies on human decision-making. Visualization strategies are supported by software tools. The R packages ggdist and tidybayes are available from the github repo (http://github/mjskay).

Stan Altan lead a roundtable discussion about the name (and hence identity) of nonclinical statisticians. Various of our colleagues reported on organizational structures for nonclinical statistics in their companies.

The best paper award is now in honor of our friend and colleague Stan Altan, a distinguished statistician with a career supporting the goals of this conference. The third-place Stan Altan Paper Award was given to Percival Sondag and Pierre Lebrun. The second-place Stan Altan Paper Award was given to Steve Novick, Elizabeth Christian, Erika Farmer, and Max Tejada. And the first-place Stan Altan Paper Award was given to Richard Burdick, Neal Thomas, and Aili Cheng.

The second poster session featured a variety of statistical topics from six different speakers.

Paul Faya discussed a Bayesian and mixed model approach to method validation, borrowing information from prior studies. Salient concerns included prior selection, in a way that regulates the relative influence of data and prior. Power priors were employed. These tools allow for the accumulation of information over time. This work was a joint effort by the DIA Bayesian Scientific Working Group and the Biopharmaceutical Nonclinical Working Group.

Vasudha Sehgal spoke about genetic and other predictors for the effectiveness of a treatment for non-small cell lung cancer. Clustering and other statistical tools were applied to this analysis, and training and test data divisions were used for model construction. Differential effectiveness based on genetic groups was demonstrated.

Matilda Hsieh applied flexible but complicated nonlinear mixed modeling to PET imaging data, and fund results whose difficulty was dependent on the organ whose data was modeled. Curve fitting and visualization were used, and random effects were employed.

Binhuan Wang applied data mining techniques to compositional data. The data of interest are represented as matrices. Investigators looked for row and column clusters. Biclustering was applied. Convexity was enforced to guarantee a global optimum. Solutions are defined by the Sylvester equation; this represents a novel application of these techniques in statistics. The method also naturally

provides for incorporation of compositional constraints. These techniques were applied to microbiome data.

Jose Ramirez and Fang Chen investigated stability shelf-life limits, using Bayesian methods to borrow information from prior studies. Power priors and mixed models were used. Heavy-tailed priors were used to model differences between drug lots. Choosing a prior distribution to match prior belief is a challenge.

Recordings for all sessions are available to replay on the conference platform. Look forward to seeing you tomorrow.

Thursday, the last day of the conference, Steve Novick opened the conference with thanks to NISS, Janssen, and Rutgers for conference support. Steve also thanked an anonymous donor. We really appreciate this support. Steve notes our sponsorship by the ASA Biopharmaceutical Section, and is led by the Nonclinical Biostatistics Leadership Forum. Steve notes that nonclinical biostatistics is broad enough to include all biopharmaceutical statistics that is not part of clinical trials. Steve also pointed out other NCB initiatives, including communications and outreach, scholarship, and workstreams on p-values and Bayesian statistics.

Dave LeBlond, Stan Altan, Paul Faya, and Don Berry presented perspectives on regulatory guidance in CMC. The consideration is cutpoint interval determination. Dave noted the imperative of sharing experience in Bayesian analyses. Stan noted late developments of regulatory guidance in clinical trials and medical devices, and notes a lack of guidance in CMC. Stan points out that Bayesian method adoption in CMC will precede guidance, and asks what CMC statisticians will be guided on. Paul noted some reasons for less guidance in CMC, recounted Bayesian guidance on clinical trials, and noted that principles driving clinical guidance could be applied to CMC. Don recounted some history of Bayesian ideas in clinical trials, related interactions with the FDA, and noted that making an impact on clinical trials was slow and hard work. Don further described recent progress, and discusses the principle of least-burdensome inference. Robust debate followed.

Roger Hoerl discussed big data impacts on statistical science and statistical engineering, and noted current large impact. He notes that some voices point to disappointment with the advances in big data in comparison to the early promise, that large data sets cannot replace critical thinking. and that the ultimate aim is generally to solve a problem rather than fit a model, and urges an engineering view of statistics, and the statistical thought process ought to be better documented.

Richard Higgs discussed statistical models to predict anti-drug antibody incidence, in order to remove threats to emerging therapies. Data were amalgamated from a variety of sources. The constructed model indicates therapies that are likely to be rejected by a large number of potential patients, and the model was interpretable rather than a black box.

Carson Sievert discussed augmenting data exploration with interactive graphics. Carson notes that interactivity adds quite a bit to the ability to perform dynamic data analyses, but requires computational

groundwork. HTML widgets provide this groundwork. Interface was exhibited using the package plotly. Interactions of various R Studio packages including HTML widgets, RShiny and plotly and their utility in applications was showcased.

Stan Altan and John Kolassa lead a roundtable discussion on p-values. We appreciate Ron Wasserstein's presence in the discussion. The discussion was lively and productive.

The first place award for best student paper goes to Louise Leonard, and the second-place award goes to Jinghang Lin.

Helena Geys reviewed the use of p values in nonclinical biostatistics. Results presented are those of a variety of authors. Helena reviewed the history of p-values, and the history of criticism of the p-value, and argues that the term "statistically significant" is problematic. Results were discussed by area (CMC, Safety/Toxicology, and Discovery). and noted that many preclinical applications are exploratory. Each of these areas has its own challenges. In many instances, confidence intervals can provide more information, and Bayesian approaches can represent an improvement.

Dwaine Banton reported on monitoring stability and early formulation selection of biotherapeutics. Of interest in this case was stability under thermal stress, using historical data. Bayesian and surrogate tools were employed. The investigation was motivated by the need for quicker formulation. Communication with scientists was very useful, particularly in focusing and prioritizing investigations. Prior elicitation was considered.

Eve Pickering discussed pre-clinical experimental rigor. Both external and internal projects have resulted in costly non-reproducible results, and points to design and the assay compatibility tool as a way to control reproducibility. Eve pointed to a cross-disciplinary task force for reproducibility.

Ken Goldberg spoke on incomplete block design for cut point detection in assays. Application is to antibody reactions to large molecule therapies. Ken pointed out drawbacks in common published designs, and in particular to the lack of randomization. Ken further presents recommended designs.

Valeria Sherina discussed workflows in R for mass spectrometry data. Data are too large and complex for local data storage and analysis. Salient questions revolved around sample size, in terms of animals and tissue sections per animal. Mixed models were fit, and supervised and unsupervised learning were also used.

Christian Schmid presented an application of dissolution testing in a continuous manufacturing process. Traditional testing is expensive and time consuming and a cheaper alternative is desirable. An approach based on DOE followed by surrogate model building was described to address this challenge. The surrogate model was used for model prediction of solution performance after a given time interval (20-30 minutes). The performance of the surrogate model was evaluated on simulated batches in Bayesian framework. Overall, this approach shows promise for the Real time release testing (RTRT).

Xin Huang closed the meeting. Xin thanked all of the attendees, and congratulated meeting committee

meeting members. Xin notes that this conference went very well and that we had good interactions during roundtable sessions. The length of time we needed for planning made the virtual aspect of the meeting necessary. Our registration numbers began below those of the last two conferences, but we eventually caught up to the last two years. Xin notes that this is a "fat tailed event". We had 109 US participants, and 17 foreign participants, from a wide variety of regions and sectors. The Pathable platform shows much communication. Session recording will be available for 1 year, and all presentations are eligible for the SBR journal. Xin recognized our conference staff. We hope to see you in 2023 in person at Rutgers.

Reminder: a post conference survey will be sent out next week.