Response to NCTR SAB Subcommittee Review of the Division of Bioinformatics and Biostatistics (DBB), on May 19-20, 2022

February 16, 2023

On behalf of DBB, I would like to express my sincere appreciation for the effort and time that the Subcommittee members spent reviewing the division. The review of so much information is not easy, and the quality of the report clearly demonstrates the devotion from the Subcommittee members, for which we are very grateful. We take the SAB Subcommittee's recommendations/suggestions/critiques seriously and they will guide our next phase of research and support activities within DBB.

The response below is a collective input from the division leadership including all leaders of the five Focus Areas. We appreciate all the positive comments such as "the SAB notes that overall, DBB's achievements have been outstanding in all major areas of its activities including basic research, collaborations, training, and support of other Centers." These positive comments offer tremendous encouragements to the division staff, and we thank you. We followed the structure of the report to prepare this response (see TOC below) which addresses the concerns/suggestions/critiques in a point-by-point format; our responses are in blue while the original text remains in black.

Sincerely,

Weida Tong. Ph.D.
Director, Division of Bioinformatics and Biostatistics
NCTR/FDA

I. "Overview" Section

Recommendations for the DBB.

- Overall, DBB is an extremely valuable resource for both FDA and the research community at large, providing important regulatory science tools including important databases, software, and expertise in bioinformatics and computational toxicology. As the importance of data science and AI are increasingly appreciated by both the agency and the external research community, and the demand for DBB expertise is likely to increase, it would be helpful for DBB to establish clearer approaches for project selection and prioritization. The current approach was not made clear to the SAB.

This is an interesting observation and a helpful suggestion. The project selection and prioritization have been traditionally carried out at the center level, which involves two sequential steps, concept paper and full protocol, both are reviewed by reviewers/scientists from our sister centers. Therefore, at the division level, the division director usually is a gatekeeper for a protocol to enter the NCTR protocol system, which is usually evaluated in accordance with the division mission. For DBB, our mission is "to ensure that the division's activities relate to FDA's review process, our linkages with product centers continue to be strengthened, and our capabilities evolve to meet the FDA's current and future needs." For that, DBB PIs are advised to follow a general guideline to develop a research project. The last paragraph of page 11 of the "SAB Review Document" summarizes this guideline, which is "All division research protocols have followed a general **ICE principle;** a division project should be **Impactful** to the FDA's mission, **Collaborative** with our sister Center(s), and **Expandable** from research to review. We specifically emphasize that all our research protocols must involve scientists/reviewers from other FDA Centers to ensure that our research has some level of impact on the FDA regulatory mission." With that said, I see the rationale behind this suggestion, and I will leave it with the senior leadership at NCTR to assess a need of center-wide implementation across the NCTR divisions.

- It is obvious and commendable that DBB has achieved prominent, and growing recognition by multiple Centers within FDA. It would be helpful to outline in greater detail how DBB collaborates with other branches within NCTR.

For the DBB initiated research projects, around 60-70% of them have collaborators from most, if not all, NCTR divisions. Equally, for the research projects initiated by PIs from other divisions, division staff serve as co-PIs to help bioinformatic and statistical analysis. These collaborations are grassroots efforts. We have also developed and maintained critical infrastructure and platforms (such as Galaxy platform for genomics) to support NCTR-wide research involving bioinformatics.

- SAB supports the efforts of the DBB to establish stronger working relationships with ORISE and Arkansas State University as it enables the Division to effectively recruit additional minds and hands to work on important and interesting projects. It would be advisable to continue in this direction and consider making similar strategic arrangements with selected academic institutions across the US.

This is an excellent suggestion, and we will continue this direction by engaging other institutions beyond Arkansas.

- It is clear that research developments by DBB have achieved prominent recognition within FDA as reflected by multiple awards received by DBB scientists. It is important and highly advisable to increase DBB visibility at the national and international levels, increase collaboration with external researchers outside of the Agency, and promote the distribution of databases and tools via specialized publication types (such as Application Notes supported by such journals as *Bioinformatics* and *Journal of Chemical Information and Modeling* and database issue supported by *Nucleic Acid Research*).

Thanks for the suggestion and we will continually build-up and strengthen this endeavor. Specifically,

- We will continue to increase DBB visibility in national and international societies by organizing both national and international conferences. In 2022, we had several activities in this area (see the list below). In 2023, we will continue to engage these activities.
    - April 2022: Led the organization of the 18th Annual Conference for the Mid-South Computational Biology and Bioinformatics Society (MCBIOS2022) on "AI and Data Science in Drug Development and Public Health"
    - May 2022: Co-organized the workshop "Cheminformatics Resources of U.S. Governmental Organizations"
    - September 2022: Co-organized the 5th Annual Conference of the MAQC Society, which is held in FDA White Oak Campus (DC) and sponsored by NCTR
    - September 2022: Co-organized the ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop
    - December 2022: Presented the R2R program at the inaugural FDA Digital Transformation Symposium
- We will continue to deliver presentations at national and international conferences. DBB senior leadership (branch chiefs and Dr. Tong) were invited to present close to 20 national and international conferences since the Subcommittee review till the end of 2022. In 2023, we will continue to improve our presence and visibility with all division staff to engage these activities, particularly the senior leadership in the division.
- We will increase collaborations with other government agencies such as NIH, universities in the U.S. and other countries, publishing application notes in the suggested journals (and others) on our developed databases and software tools, etc.

- The report provides information about cumulative successes of the Division and average productivity by DBB staff. For the purposes of fair productivity assessment and career advancement, it would be helpful to outline metrics by which DBB staff are assessed individually.

The FDA has a standard mechanism to assess performance of a federal employee, which is called the "Performance Management Appraisal Program" or PMAP for short. In this mechanism, a supervisor establishes a metric which is first communicated with his/her subordinate at the beginning of the year, followed with a mid-year progress review, and a final scoring of performance at the end of the year. DBB established the metrics for both research and support scientists tailored to their grade levels (pay schedule) with general instruction on the steps each individual need to take to move up the grade.

- The Division supports multiple projects important for the field of regulatory toxicology, including the development of multiple computational toxicity prediction models. These models are

disseminated via research publications and presentations at scientific meetings. Given the overall DBB and NCTR missions, it would be important to outline specific steps toward making such models into accepted regulatory tools that are used routinely by the agency.

We have three different pathways to make our models useful in regulatory application:

1. Consultation: we have provided consultations to assess the liver toxicity with our models in CDER review.
2. Deployment: We are working with CDER to deploy our model in a secure environment where the regulatory data can be directly input into our model for proof-of-concept.
3. Qualification: We worked with the CDER Innovative Science and Technology Approaches for New Drugs (ISTAND) to qualify our models.

- Because of its demonstrated and well-supported value to the development of regulatory science at FDA, DBB is encouraged to continue its outreach to other Centers, Offices and Divisions within FDA. The work done to date is potentially extensible to other areas that could open new areas of collaboration that contribute to both DBB's and FDA's mission.

We appreciate the suggestion, and we will continue to build-up and strengthen this effort.

II. "Focus Area 1: Regulatory Applications and Support" Section

Overall, DBB has performed a remarkable amount of work on behalf of FDA Centers in a relatively short period of time. Notably, DBB has developed a number of tools to integrate IT and machine-learning/assisted solutions to allow FDA to better search, capture and assess large volumes of information that previously could only be managed through laborious free-text searching and manual capture and maintenance. These tools have demonstrated a high degree of user acceptance and rapid uptake by FDA staff within the Centers showing the importance of DBB's work in strengthening the review process and building efficiency. Given the success of these programs, DBB is urged to proactively find ways to expand this work across more work areas within the Centers to continue to gain IT-enabled efficiencies and transparency within regulatory science and review support. Expansion of these tools and closer integration with other FDA Offices and Divisions would only serve to make the Regulatory Application and Support Focus Area stronger and even more relevant to the FDA's and NCTR's missions.

DBB shares the SAB Subcommittee's desire to expand the reach of its Regulatory Application and Support efforts. Since the site visit in May 2022, we have developed additional Smart Templates for the CDER Office of Clinical Pharmacology and have had exploratory discussions with the CFSAN Office of Food Additive Safety. We intend to seek additional collaborations with FDA Centers, Offices and Divisions.

III. "Focus Area 2: Alternative Methods and Knowledge Bases" Section

The first area, the efforts to develop KBs has been productive. Five KBs have been developed and are under refinement: Estrogenic activity database, Endocrine Disruptor KB (EDKB), Liver Toxicity KB (LTKB), NCTR liver cancer database and Tobacco constituents KB. LTKB is a publicly available, mature KB of drugs associated with DILI. DILI continues to be a human health issue as well as a stumbling block for development of new pharmaceuticals. LTKB has been a useful resource for a variety of audiences, including researchers. The Division is encouraged to continue their efforts to incorporate additional factors, including immune and genetic factors, into the database. A relatively recent effort is the development of a database of herbal dietary supplements (HDS) associated with cases of liver injury. DBB is using PharmaPendium, an Elsevier database of historical agency approval documents, to incorporate pharmacokinetic (PK) and adverse effect data to develop predictive models for DILI. This will assist FDA reviewers and scientists. Through a recently approved proposal with the Office of Women's Health, DBB will use AI-based technologies to explore immune system-related risk factors to help explain sex differences (women represent most cases) in HDS-induced liver injury. DBB is encouraged to expand their efforts beyond immune factors.

We thank the Subcommittee for recognizing the usefulness of the developed knowledge bases and databases and for the great comments to improve our work in the future. We will incorporate additional types of data related to liver toxicity into the liver toxicity knowledge base (LTKB). Specifically, we plan to extract the immune system-related and genetic factors that are associated with liver toxicity and incorporate them into the LTKB to enhance its data scope. The Subcommittee's recognition of usefulness of the developed herbal dietary supplements (HDS) database to FDA reviewers and scientists and encouragement to expand our efforts to the current database is greatly appreciated. We will mine genetic effects and mixture related to HDS from public sources including public databases and literature for improving diversity of the data in the database.

Also under development is a KB of chemicals and their associated opioid agonist/antagonist activity (OAK) using data generated in house and curated from the public domain as well as other descriptive data. This effort addresses the health crisis surrounding opioid abuse. The goal is for the OAK to provide insights to molecular mechanisms of pain management and treatment of opioid use disorder and also to enable development of *in silico* models for predicting opioid activity. Data points have been collected for over 2800 compounds, and ML algorithms are being used for models to predict activity. This is an important effort that needs further refinement and thought concerning how it could be used to facilitate treatment of pain and opioid use disorder.

The Subcommittee's suggestion to further refine OAK is well accepted. For that, we will augment opioid agonist/antagonist activity data and investigate the various QC metrics for inclusion/exclusion to ensure the quality of the database. The improved quality and quantity of the data in OAK will be used in the development of machine learning models to further enhance performance for predicting agonist/antagonist activity profiles of drug candidates, facilitating the development of treatment for pain and opioid use disorder.

DBB is initiating efforts to develop a database of molecules with androgenic activity as a resource for FDA scientists and reviewers. There was no mention of trying to identify motifs within chemical structures that confer activity or build computational models with this data (similar to proposed developments for the OAK database), so the team should be encouraged to pursue this avenue.

We thank the Subcommittee for raising the question on missing information on identification of structural motifs and development of computational models in the presentation of our efforts to develop a database of molecules with androgenic activity. We focused on the data curation and database development in the review meeting as they were ongoing at that time. We agree with Subcommittee's suggestion and will develop predictive models using various machine learning and deep learning algorithms for predicting androgenic activity of chemicals. In addition, we are planning to identify key structural motifs that are related to androgenic activity of chemicals through analyzing frequency data of chemical features that are used in machine learning and deep learning models.

With respect to the second research area, development of predictive models for DILI, DBB is using the Rule-of-2 and a DILI score to identify inherent hepatotoxic characteristics of drugs under review. They plan to incorporate PK data from PharmaPendium and employ a ML approach to develop a predictive model that will then be validated with failed drug candidates. They also plan to use PBPK modeling to predict some additional PK parameters. The approach is reasonable, although it was not made clear how these models would be used in the review process.

We appreciate the Subcommittee for recognizing our approach to incorporate PK data into the development of machine learning models for predicting drug induced liver injury (DILI). We appreciate the opportunity to clarify utilization of the models incorporating PK prediction and PK data in the FDA review process. As summarized in page 3 that we have three pathways to support the review process with our models; these are consultation, deployment, and qualification. Similar to the RO2 model, we would likely use this model in the consultation process.

One project was presented related to the third research focus area, developing predictive models as potential alternative methods to animal toxicity testing. This project is in the early stages and, as a pilot, data from guideline animal studies of multigenerational reproductive toxicology were used. The model was reported to have achieved "reasonable predictive power" although accuracy was 60-65%, raising a question of what is considered reasonable. Accuracy can be improved with increasing confidence in a prediction confidence analysis. Future studies will involve developing several DL and ML models, expanding to other toxicity types, and exploring modeling approaches to improve predictive accuracy. It was not entirely clear how the model is intended to be used, but if it is to be used as an alternative to animal testing, the criteria for applicability need to be defined.

We appreciate the Subcommittee for raising the question about "reasonable predictive power". Because animal toxicity is a complex mechanism and involves many factors, animal toxicity is usually more difficult to predict than in vitro toxicity. The qualitive conclusion "reasonable predictive power" is based on our previous experience on other types of animal toxicity, it is not a quantitative one. We will improve our models by increasing the amount of data used for training and including mechanism-related information.

The improved models will serve as a screening tool to identify areas of concern for follow-up experimental validation. In other words, our model is intended assist the animal studies, but not for replacement. We agree with the Subcommittee that the criteria for applicability should be defined. We are considering the OECD guideline to define the applicability domain for every model developed in our group.

With respect to the fourth research focus - using AI for alternative approaches in toxicology - one project was presented comparing ML and DL. The hypothesis was that ML would produce better results with small- to moderate-sized datasets whereas DL would be better for large datasets. It was stated that ML models for predicting CH_4 and CO_2 adsorption capacities of metal organic frameworks were high performing. These datasets are large, and DL models were constructed to compare performance. DL improved performance only marginally. One question that was raised was how large versus small is defined for datasets. The DBB has plans to decrease sample size to evaluate when ML outperforms DL. Other issues raised by the Subcommittee that should be addressed by DBB related to the availability of datasets large enough to require DL, and the criteria to assess improvement using DL instead of ML. There is a natural interest in the computational research community as to when DL techniques should be used to develop better models and when their application is not warranted as the standard ML approaches will provide models of the same or better accuracy as DL-based models. This is an interesting line of investigation that should be pursued further.

We thank the Subcommittee for raising these questions. The first question is about how large versus small was defined for a dataset. We assume that the question is in the context of comparing conventional ML against deep learning (DL) methods, as we plan to do in this project. A model's performance is usually governed by three major factors; sample size, signal/noise ratio in both class labels and the dataset itself, and data prevalence. We speculate that these factors also affect the choice of ML over DL, or vice versa, in an interactive way. Among these factors the sample size does play a major role. It might be difficult to quantify "large" or "small", which is often subjective and not quantitative. Our project focused on examining several datasets by varying data size on the comparative performance between ML and DL. Hopefully, this exercise will provide some preliminary findings to shed light on the question of "large versus small" in a semi-quantitative way and might also guide us in future research on defining "the availability of datasets large enough to require DL, and the criteria to assess improvement using DL instead of ML." We agree that there is an immense interest in the computational research community to know when conventional ML or DL should be applied since the latter tends to be difficult to explain. Our effort is to conduct comprehensive comparative investigation using the commonly used ML and DL algorithms based on various sizes of datasets to provide the computational community some results that are helpful in selecting appropriate ML and DL algorithms for datasets of interest in the future.

Summary and additional observations.

Overall, work presented under these focus areas has been very intense providing an important contribution to the DBB mission. Further research developments are encouraged especially concerning the expansion of curated knowledge bases and making these accessible to the research community for building/benchmarking respective predictive models for respective endpoints. The work on making models both accessible (continuous sharing via github) should be encouraged and even more so, it is

important to emphasize the use of models as NAMs achieving acceptable regulatory tool status, for instance, via registration for the MDDT program.

We really appreciate the Subcommittee for these constructive suggestions. We will continue to expand our research efforts in curating data, developing knowledge bases, and making the data and knowledge bases accessible to the research community, in addition to the FDA reviewers, for building/benchmarking predictive models for respective endpoints. We accept the Subcommittee's suggestion to share our research results via github and registration for the MDDT program for selected models.

IV. "Focus Area 3: Precision Medicine and Therapeutics" Section

Scientific projects:

Long-standing efforts within the Division, and the first area of interest have focused on the international consortium of Microarray and Sequencing Quality Control (MAQC/SEQC). This consortium was established to address the reliability, reproducibility, and utility of genomics technologies in regulatory use and clinical application, and application of drug repositioning principles to assessing reuse of existing drugs for the treatment of rare diseases and COVID-19. … The success of this program is evidenced by the number and quality of published works in high quality journals such as Nature Biotechnology and Genome Biology. Proposed plans to engage with PrecisionFDA to further evaluate analytical tools for indel calling following Oncopanel sequencing is of major relevance in precision oncology. However, the degree to which this effort is being coordinated in collaboration with academic stakeholders and other government agencies should be considered.

This is an excellent suggestion. While such an arrangement could be challenging under the PrecisionFDA umbrella, we will explore the opportunity with the MAQC International Society (https://themaqc.org/) which is derived from the MAQC consortium. The society has an established mechanism to facilitate the collaboration between government, industry, and academia.

A fourth area of focus is drug repositioning for rare diseases taking advantage of AI-powered frameworks. … It is not clear if these will be activities listed to bring to closure this area of focus or if plans are in place to continue these efforts into the future. As noted in the previous review, it is also not clear how this overall effort fits into the mission of NCTR and FDA.

Thank you for the question. The PI leading this project departed DBB/NCTR at the end of 2022 and his replacement had not yet been identified. Consequently, this effort has been postponed.

We do apologize for not clearly demonstrating the relevance of this program to the FDA mission. FDA has an active program on orphan drug development and rare diseases. The program is governed by the Orphan Drug Act which is a law passed by Congress in 1983; the program incentivizes the development of drugs to treat rare diseases. For that, the Orphan Drug Act established the Orphan Product Grants Program to provide funding for developing products for rare diseases or conditions, which is administrated through the FDA Office of Orphan Product Development. Consequently, regulatory science in this area provides the scientific knowledge necessary to facilitate and regulate orphan drug development. In 2022 we were invited to present our program during FDA's Rare Disease Day (https://www.fda.gov/industry/fdas-rare-disease-day/fda-rare-disease-day-2022). This is an annual event organized by the FDA Office of Orphan Product Development. This year's meeting will be held February 27, 2023 (https://www.fda.gov/news-events/fda-meetings-conferences-and-workshops/public-meeting-fda-rare-disease-day-2023-02272023#:~:text=Event%20Title,February%2027%2C%202023) and NCTR was again invited to speak.

V. "Focus Area 4: Artificial Intelligence (AI) and Machine Learning (ML)" Section

The first area presented by DBB was on AI-Based NLP for FDA Labeling Documents. …Similar effort relying on NLP is also ongoing in the second area of research on DeepReview, an NLP Powered Information Retrieval System. Of a minor comment, dividing the NLP-based developments into two separate areas appears somewhat artificial so it would be advisable to integrate these two projects under the same NLP umbrella.

We completely agree with the SAB Subcommittee that these two and other NLP-based projects could be integrated under the same NLP umbrella. One of the advantages for an NCTR team, instead of a contractor, to implement AI techniques for the Agency's regulatory operations is that the underlying technologies and components can be shared across developed projects to advance all projects. We continue to actively pursue cross project sharing and integration.

A third area covered during this SAB meeting was on the development of SafetAI, which is a CDER initiative to enhance the IND review process. This area is a continuation of the traditional expertise of the team in developing QSAR models for important toxicity endpoints but now with DL methods. The team is also proposing to enrich current models that have been built historically with chemical descriptors by developing a new modeling framework where instead of chemical descriptors, the investigators will use "model-level representation from conventional ML classifiers into a DL framework. This description is somewhat unclear; hopefully future studies and publications will help establish both the utility of these models and their advantages (of any) as compared with traditional ML techniques.

We very much appreciate the suggestions and do apologize for not describing clearly this in-house method. Basically, this is a DL version of consensus approach where the prediction results from multiple models are integrated with a deep leaning architecture. The framework consists of base classifiers and a meta classifier, where the outputs from base classifiers (model-level representation) are the input for the meta classifier which consists of a three-layer neural network. In a paper we are preparing for mutagenicity prediction with this method, we compared it against five common machine learning methods. At the time of preparing this report, the paper is undergoing NCTR internal review, and we will provide an update at the upcoming SAB meeting in April.

The fourth area represents an effort toward achieving eXplainable AI (XAI) in Regulatory Science. This is an interesting and important area of investigation, especially in regulatory science where transparency of models, especially statistical models, is always preferred. Deep learning approaches that are widely employed in AI methods, are notorious for the lack of transparency and clear interpretability so the effort toward explainable AI is understandable. While the importance of this direction of research is obvious, the description of the proposed effort does not seem to address the complexity of DL model interpretation in terms of feature significance. Preliminary observations presented by DBB suggest that simple models (such as linear regression) may offer the best balance between accuracy and explainability. It appears that it will be important to address the utility of DL models for various endpoints as compared with simpler models and then contemplate DL model explainability only for those cases when the use of DL models is justified.

We appreciate the Subcommittee's recognition of the importance of this direction. We apologize for not articulating clearly regarding the approach we are taking to address the complexity of DL model interpretation in terms of feature significance. We have published two papers on this topic, both of which applied do-calculus on text documents to establish the causal relationship of feature significance and model performance (see the reference below). Currently, we are evaluating this methodology on every DL method we are implementing. In addition, we have also developed new explainability metrics for general use in ML and DL. AI explainability is an active field of research. We will closely monitor the research field and leverage any progress that may emerge in the community. As an update, our recent proposal won the highly competitive Challenge Grant from FDA's Office of the Chief Scientist to assess and mitigate bias in applying AI-based NLP to drug labeling documents. This will become an integral part of our AI/ML research portfolio. The work on bias evaluation and mitigation for AI models is also closely related to AI explainability. Making AI models increasingly more explainable is key to correcting the factors which inadvertently lead to bias.

- **InferBERT: A Transformer-Based Causal Inference Framework for Enhancing Pharmacovigilance**.
  Wang X., Xu X., Tong W., et al.
  Frontiers in Artificial Intelligence. 2021, 4:659622-659622.
- **DeepCausality: A General AI-Powered Causal Inference Framework for Free Text: A Case Study of LiverTox**.
  Wang X., Xu X., Tong W., Liu Q., and Liu Z.
  Frontiers in Artificial Intelligence. 2022, 5:999289.

Summary and additional observations

Overall, this area of research continues to be one of the major components of the DBB research portfolio. Importantly, the Division continues to stay at the forefront of using modern data analytical methods as applied to problems in computational and regulatory toxicology. Several novel initiatives relying on advanced NLP and DL approaches are noteworthy. It will be important in the next few years to establish the relative value of these new approaches as compared to more conventional ML methods as well as achieve the broad use of novel document review tools across the agency. It will be also important to continue to promote databases and tools created by the DBB across research community at large.

We appreciate the advice for a sharp focus in the next few years to guide this area of research in AI/ML.

- For a new method explored and developed in the division, a comparative analysis against the standard ML methods will be carried out with respect to statistical performance, explainability, context-of-use, adaptability, and regulatory implementation. We are preparing a manuscript which examined all of these aspects for our in-house new DL method with a large dataset of >13,000 compounds. We are also working with CDER to implement our model in the regulatory environment.
- With regard to applying document review tools with AI across the agency, we have established an international working group to work with other regulatory agencies to explore and establish guidelines for applying NLPs to regulatory decision-making.

- We agree that we need to continue making the community aware of our efforts and products in this area. We have a wonderful communication team at NCTR, and they have set up several webpages about our tools which can be found by Googling “FDA Bioinformatics Tools” or “AI4Tox”.

VI. "Focus Area 5: Real-World Data (RWD) And Real-World Evidence (RWE)" Section

The Biostatistics branch currently has eight employees (one vacancy), including six research scientists, Drs. Wang, Baitang Ning, Paul Rogers, Wei Zhuang, Zhiyuan Lu, and Dongying Li. The role of the research scientists in this branch focuses on research rather than supporting other studies carried out at NCTR, as noted in the Division's booklet. This is a change compared to the 2015 review of this branch. It is unclear whether this is a shift of focus for this branch, and if so, what the underlying motivation was.

This is an excellent observation. There has been a significant turnover in the division since the last review in 2015 and the Biostatistics Branch has been most affected. Specifically, due to a center level reorganization in 2019, NCTR has moved most of the statistical support from other NCTR divisions to the Office of Scientific Coordination. Therefore, the Biostatistics Branch focuses on methodological research. In addition, the branch dealt with the retirement of two Chiefs prior to the appointment of Dr. Wang, and all the members in this branch are relatively new. This drastic change in both staff and leadership poses a challenge in terms of continuity but also provides an opportunity to assess the future direction of this branch. Under Dr. Wang's leadership, this branch now focuses on research of FDA's emerging needs. More specifically, the branch has been very active in research involving RWD/RWE, which is a significant area of focus in FDA. This has been carried out in a collaborative fashion, across DBB branches, with scientists from other NCTR divisions, and other FDA centers (e.g., CDER).

The first area presented at the review by Dr. Paul Rogers focused on developing a Charlson Comorbidity Index (CCI) for American Indians, mCCI-AI. The motivation of the project was the lack of such an index for American Indians. The construction of mCCI-AI is an extension of Charlson's method which was built upon a Cox proportional hazards model with random family effects addressed. Although the Subcommittee agreed that there was a potential need for such an index, compared to CCI, the advantage of mCCI-AI was not clearly emphasized and the quality of the newly developed index, mCCI-AI, was not discussed.

We do apologize for not clearly articulating the advantage of mCC-AI. As the Subcommittee is fully aware, the Charlson Comorbidity Index is the most widely-used score to measure comorbidity when the outcome is mortality. The development of the mCCI-A is the first of its kind for Native Americans. The mCCI-AI can distinguish between subjects that died and those that remained living within one year of their Phase VI of the Strong Heart Study (SHS) exam 73% of the time. Conversely, the original CCI only achieved 66% for the same Native American population.

The second area presented by Dr. Wen Zou focused on prescription opioid-use-associated safety signals. The project had several stages starting from a systematic literature review, followed by data retrieval from different sources and data mining and deep learning. Findings from a systematic review was presented at the 2021 IEEE International Conference. The projects are of interest and relevant to FDA missions. It will be interesting to see whether the findings from the retrieved data are consistent with findings in the literature.

We appreciate the comments from the reviewers. Recently, we had new findings in the retrieved data and identified many points of consistency between the retrieved data and the literature. The results from this work were summarized in a new manuscript

The third area was delivered via a poster by Dr. Dongying Li. It is a bioinformatics project focusing on miRNA biomarker detection. This research program addresses an important area and has made significant progress in identifying important leads for future development. The ability of the team to identify specific targets for future development will be critical to ensure continued success of these efforts.

We appreciate the reviewer's comment. This project follows the FDA's Predictive Toxicology Roadmap and aims to improve risk assessment in drug development and review. The multi-dimensional dataset and methodology enhances the significance and utility of our investigation outcome.

The fourth area utilized propensity scores to assess the impact of racial/ethnic factors on critical care delivery. Using propensity score to match has a potential of overfitting. It will be informative to see the robustness of the propensity scores inferred in the study in terms of matching/weighting and the final conclusion drawn for associations between ethnicity and critical care delivery.

We agree that utilizing powerful approaches such as propensity scores requires extra care. To ensure that propensity score models do not introduce unwanted biases in inference, we will carefully perform validation for the proposed propensity score models. We also plan to carry out sensitivity analysis to assess the robustness of our conclusions by artificially introducing extra bias in the data and evaluating the impact on the results. This will provide assurance that our conclusion reflects the salient patterns in the data.

Final notes.

It is noted that the budget provided by FDA to DBB was significantly reduced in year 2021 ($500K in 2021 compared to >$1milion in 2020) and the reason for this reduction was not discussed. However, compared to the year of 2015, the number of staff members appears to have increased (to ~60 from ~50). New focus areas have been added to the Division to fit FDA's new needs. Also, a large amount of data including high dimension high throughput omics data require AI-techniques to process. All these seem to support the necessity of recruiting more manpower to accomplish all objectives and challenges faced by DBB. Related to the need of manpower, we also noticed that there are a number of vacancies in DBB (currently, 5 vacancies not including the Scientific Computing Branch). The Division has taken actions to ease this condition, e.g., recruiting from local universities through various programs. On the other hand, a stronger retention plan from FDA as an institute will certainly help the Division maintain a strong and productive team.

Thank you for the inquiry about the low budget we had in 2021. We had a bad year in 2021 with a significant reduction in the NCTR discretionary budget for various reasons.

At the time of preparing this document, we have successfully filled 3 out of 5 vacancies by converting on-site postdoctoral fellows to staff fellows. With that said, there is extreme competition in this rapid moving area and we continue to lose staff to industry. For example, Dr. Zhichao Liu (Technical Lead of the AIRForce team) just took a job in a pharmaceutical company. This is a government-wide issue as we cannot offer the higher salaries they can obtain in many industries, and this disadvantage is especially pronounced for NCTR.