# NCTR Division of Systems Biology Review

## November 3-4, 2016

## Background and Overview

The Division of Systems Biology (DSB) is one of six Divisions at the National Center for Toxicological Research (NCTR). This Division formed in 2004 has undergone multiple organizational changes, the most recent being splitting off the Division of Bioinformatics and Biostatistics in 2012. The DSB consists of three (3) administrative staff members, seven (7) post-doctoral fellows, eleven (11) support scientists, and twenty-three (23) research scientists. There are four main scientific thematic areas namely,

a. Translational Safety Biomarkers and Mechanisms
b. Alternative Drug Safety Models
c. Food Safety Technology
d. Computational Modeling

Dr. Daniel Acosta (Deputy Director of Research, NCTR) and Dr. Donna Mendrick, Designated Federal Official (DFO) and Associate Director for Regulatory Affairs at the FDA/NCTR in consultation with the NCTR leadership and DSB organized the review team and provided the "charge" to the review team. The review team comprised individuals from academia and private industry with expertise in microbiology, toxicology, cardiology, and pharmacology (Table 1).

Table 1: NCTR Division of Systems Biology Review Team members

| Name | Affiliations | Role |
|---|---|---|
| Suresh Pillai | Texas A&M University/Member, NCTR Science Advisory Board | Co-chair |
| Diwakar Jain | Westchester Medical Center/Member, NCTR Science Advisory Board | Co-Chair |
| Nigel Greene | AstraZeneca | Member |
| Frank Barile | St. John's University | Member |
| Tim Ryan | Sano Informed Prescribing | Member |

The review team's task was providing objective advice to the NCTR Director, the DSB Division Director and the research scientists on

1. The quality of the research being conducted within the DSB
2. How the DSB could improve horizon scanning for emerging technologies and new safety assessment approaches.
3. The critical emerging regulatory, research, scientific issues, trends, and needs *vis-a-vis* the NCTR/FDA.

The review took place Nov 3-4, 2016 in Little Rock, Arkansas immediately following the NCTR Science Advisory Board Meetings. Dr. Greg Lanza, member for the NCTR SAB also joined the DSB review and contributed to the discussions.

## Agenda and Review Process

In preparation for the review, the review team received the agenda (Appendix 1) as well as background information about the DSB. The background materials included descriptions of the on-going research projects, CVs of the lead research scientists, abstracts of the scientists' presentations, copies of PowerPoint™ presentations, posters as well as a copy of the 2009 review (and the response) of the Division of Systems Toxicology. The scientists also posed a series of "open questions" for the review team. In preparation for the review, the review team held a conference call, October 26, 2016 to discuss the pending review and to develop a game plan for the review. Dr. Mendrick joined the conference call and provided her insight and perspectives into what the sub-committee review expects to accomplish. The Division presented current and future research ideas so that the committee's advice is useful for strategic planning purposes. The review team created primary and secondary reviewer assignments to facilitate review and report preparation. Based on the materials provided to the review team and the reviewers' expertise, the review team assigned primary and secondary reviewers to specific thematic areas (Table 2).

Table 2. Reviewers for the different thematic areas

| Thematic area | Primary Reviewer | Secondary Reviewer |
|---|---|---|
| Clinical/translational metabolomics and proteomics | Tim Ryan | Nigel Greene |
| Dox/Cardiotoxicity markers | Diwakar Jain | Frank Barile |
| TKI Markers and mechanisms | Diwakar Jain | Frank Barile |
| Alternative drug safety models | Frank Barile | Tim Ryan |
| Food Safety Technology | Suresh Pillai | Nigel Greene |
| Computational modeling | Nigel Greene | |
| Precision Medicine | Tim Ryan | Diwakar Jain |

## Executive Summary

The review of the NCTR Division of Systems Biology (DSB) took place November 3-4, 2016. The review team members were from academia, pharma industry, and the medical profession. The DSB has organized itself into three research branches and the Division focuses heavily on developing and evaluating biomarkers for hepatoxicity, cardiotoxicity, and studying the mechanistic aspects of organ, tissue and cellular drug toxicity. Though the DSB is made up of highly competent scientists, the physical separation of the different DSB labs could be creating in operational inefficiencies. Also, there is a compelling need for integration of systems biology with computational technologies so that there is efficient logical progression from discovery to clinical value. There was significant expertise in the DSB in identifying biomarkers There is a need for more specific (proteomic/metabolomic/genomic) biomarkers especially for cardiotoxicity and therefore the DSB's efforts in these areas are of high value. However, biomarker validation should be of high priority. Biomarker validation has strategic value to ensure that these biomarkers are evaluated at the clinical stage. The finding that there are gender-based differences in doxorubicin cardiotoxicity is extremely important observation and needs further study. The studies attempting to understand the extent of organ -specific differences in terms of tyrosine kinase inhibitor toxicity is also critically important. So also are the projects attempting to understand the relationship between obesity and drug pharmacokinetics, drug action, and drug toxicity. There was some disagreement among the review team members about the relevance and value of specific in vitro drug safety models in use. However, the team felt that these questions require further investigation and evaluation. The review team also suggested that the DSB use the recent Academy of Science report on animal testing as a road map for the development of alternate models. The efforts within DSB to develop and deploy pathogen detection and screening technologies for the food industry are noteworthy, especially their efforts at commercialization of these technologies. The research being performed in the computational modeling group is of high caliber and has the potential to make major impacts leading to advances in QSAR modeling predictions. In the area of precision medicine, the review team felt that the DSB can become the lead FDA group to lead this area of research. Overall, the NCTR Division of Systems Biology's research expertise and activities are of very high caliber and directly aligned with FDA's mission. The science at the DSB is at the leading edge of the different thematic areas and the work products from this division will have a major impact on the FDA's mission.

## Organizational Structure of the Division of Systems Biology

With the formation of the NCTR Division of Bioinformatics and Biostatistics in May 2012, the DSB has organized itself into three research branches, each led by a Branch Chief. The three research branches are the Biomarker and Alternative Models Branch (headed by Branch Chief, Dr. Richard Beger), the Personalized Medicine Branch (headed by Branch Chief, Dr. James Fuscoe), and the Innovative Safety and Technology Branch (headed by Branch Chief (Acting), Dr. William Mattes). The physical separation of the DSB labs (spread out over seven buildings) may be resulting in inefficiencies.

Dr. Bill Mattes, Director of the DSB provided the introductory overview. The mission of the Division is to address contemporary challenges in food, drug and medical product safety using a combination of systems biology and innovative technologies. The division focuses heavily on developing and evaluating biomarkers for hepatotoxicity, cardiotoxicity, studying the mechanistic aspects of organ, tissue and cellular drug toxicity. The present organizational structure maybe falling short in capturing the scientific potential of its staff. In today's big data-driven world, there is a compelling need for the integration of systems biology with computational technologies. Like many research organizations, the DSB must continually strive to integrate its diverse research capabilities to follow the logical progression from discovery to the clinic.

## Theme 1: Clinical/Translational Metabolomics and Proteomics

There was one overview presentation and three topic-specific presentations in this theme. Dr. Berger presented a metabolomic and proteomic overview and a scientific discourse on sample and data quality. Dr. Schnackenberg presented on the topics of pervasive developmental disorders and metabolomics. In the overview presentation, Dr. Berger clearly indicated that the group's focus was technology-based novel biomarker identification using well-characterized models of organ damage namely, acetaminophen- and doxorubicin-induced toxicity. Early proteomic and metabolomic studies revealed that acyl carnitines and lysophospholipids were the chosen candidate biomarkers moving forward. There is ample literature supporting acyl lipids as fluid-accessible biomarkers. The primary questions moving forward pertain to candidate biomarker validation; how will the division address precision, accuracy, specificity, sensitivity, and temporal relationships of candidate biomarkers relative to organ damage? Based upon the presentation, it was not entirely clear exactly who would demonstrate preclinical utility or translate these candidates into clinically valid biomarkers. It was evident that proteomic and metabolomic technologies are being used deployed on clinical samples originating from NCTR collaborators and the NCTR team in a sound manner. The three general questions posed to the review team were:

1) Should there be a validation of the candidate biomarkers – Answer: Yes, with experiments to address validation tenets mentioned above
2) Should candidate biomarkers be integrated into multiple technologies? Answer; Yes, using the most straightforward methodology available to demonstrate utility. New technologies that better scale or reduce cost should come with demand
3) Should NCTR get involved in novel biomarker discovery? Answer: Yes. The infrastructure is in place at NCTR to continue discovery, but be careful not to become a technology hammer looking for a nail (see comments below). Keep discovery based upon needs aligned with the NCTR mission.

Dr. Schnackenberg's presentation detailed the metabolomics and Somascan-based proteomics using aptamer capture as a means of biomarker discovery. Dr. Schnackenberg, with deep expertise is ideally suited for the technologies utilized. The studies are new, therefore, candidate biomarkers remain unidentified. The approaches presented were reasonable and should be pursued. However, the review team does suggest that the group characterize the data sets in hand prior to examining other time-points (for example, trimester samples) or moving on to lipidomics biomarker discovery. It simply makes sense to assess the information generated to date prior to moving forward. It appears that a single laboratory is responsible for generating, analyzing, and

interpreting discovery data, and as mentioned above with Dr. Berger's presentation, it was not clear if the sponsor or NCTR lab would perform validation experiments, and if so, what that validation path would look like. Using *ex vivo* MALDI imaging, Dr. Schnackenberg perform has performed mechanistic tissue distribution studies of small and large molecules for a multitude of NCTR and non-NCTR collaborators. MALDI imaging is a highly specialized and powerful technique for mechanistic study of tissue via imaging, and these studies should continue.

Based on these presentations, the review team sees a high level of expertise in identifying biomarkers in the DSB, and the laboratories have developed into highly capable specialty groups. Moving forward, the concern is that these same laboratories are technically hyper-focused, and could potentially become commoditized as technology providers offering these services in a fee-for-service model. The larger question pertains to whether the DSB has the expertise and strategic position with collaborators to validate the candidate biomarkers that they have identified. The review team feels that validation of candidate biomarkers should be a clear strategic focus of this group, recognizing that translation could involve early biomarker qualification steps with additional 'early adopting' preclinical and clinical collaborators. Focusing solely on discovering candidate biomarkers could quickly lead to unfulfilled overall clinical research goals if translatability is not clearly articulated in the strategic plan. Dr. Mattes stated the goal of bridging the non-clinical into the clinical setting, and validation packages needed to accomplish this task. Does the group have the required expertise to translate biomarkers? What evidence is there that these current biomarker candidates are translatable given NCTR data relative to the literature? What are the studies, measurements, and criteria necessary for validation? Will the responsibility of validation go back to study sponsors? The DSB should strategically address these questions. Experience shows that a team capable of carrying out all the different steps and not just candidate biomarker identification is necessary to translate biomarker information into medically relevant information. It is clear that the team is technically very sound and capable, with upfront biomarker identification technologies in place. In the case of MALDI imaging, the group could quickly become field leading. Further, the skill sets exist within the DSB to validate discovery findings. However, it is just unclear how this next phase will be executed, and whether the "long road to validation" has been strategically mapped. Dr. Berger wrapped up the session with a presentation on sample and data quality. The NCTR is a participating member in the data quality task group, and has performed internal studies on sample quality relative to storage and processing. These variables clearly contribute to the interpretation of proteomic and metabolomic data, and with the unique purview of the FDA, are necessary to objectively review data from disparate submission sources. Further, these activities support the overall FDA mission, especially as it pertains to data submission. Given the large investment that the NCTR has made in both metabolomics and proteomic technologies, a substantial integration effort should take place. This effort should include the longitudinal integration that follows the logical path of demonstrating the clinical utility of model-discovered biomarkers in addition to the integration of multiple biomarker discovery technologies presented by these research teams. The review team feels that the proven capability for biomarker discovery using model systems and the quality of the scientists and their connections with the external community to be strong points. The review team feels that it is prudent to develop a strategic plan for biomarker validation and to temper the urge to identify more candidates rather than closing out the existing work.

## Theme 2: Dox/Cardiotoxicity markers

Dr Varsha Desai and her team (Personalized Medicine) are leading the efforts to study anthracycline cardiotoxicity and to develop biomarkers for its identification. There were two presentations by Dr Desai on this aspect. Dr Desai first presented data on her mouse model of cardiotoxicity and candidate genomic, proteomic and metabolomics markers of doxorubicin cardiotoxicity. She has also conducted studies to investigate the mitigating role of Dexrazoxane on these changes. The review team was queried about the translational potential of early biomarkers identified during preclinical studies as well as the possibility of gender-based differences in susceptibility to cardiotoxicity from doxorubicin. In clinical practice a number of biomarkers of myocardial injury such as CK-MB, troponin-T, troponin-I, BNP correlate with myocardial injury associated with therapy with anthracyclines. However, none of them is specific for anthracyclines-induced cardiotoxicity. They all arise following myocardial injury induced by any insult such as ischemia, inflammation or trauma. Although, the extent of acute injury does correlate with the peak level of these biomarkers, yet, there is no good correlation between the long-term impairment of cardiac function and the levels of these biomarkers in blood. There is definitely a need for more specific proteomic, metabolomics or genomic biomarkers, based on the biochemical, molecular and cellular injury specific to the anthracyclines. The research by this team is highly relevant and needs to be extended to humans to identify genomic, proteomic and metabolomics markers to predict cardiotoxicity in clinical context.

The observation of gender-based differences in doxorubicin cardiotoxicity observed in the animal model studies is an extremely important observation. Firstly, one needs to figure out whether this is truly, gender based and not due to differences in body weights, body surface area; heart size and heart weights of the animals. If one finds different gender based susceptibility to cardiotoxicity to persist despite correcting for these factors, one needs to determine the degree of susceptibility: such as two-fold or 1.5-fold (or whatever is appropriate) increase in cardiotoxicity in female animals compared to male animals. Are the differences observable only in mice or across other animal species: rats, rabbits or guinea pigs as well? In humans, the female gender is not a widely recognized risk factor for anthracycline cardiotoxicity in the clinical literature. This is unlike other well-recognized risk factors such as extremes of ages, concomitant administration of radiation, other cancer therapies such as high dose cyclophosphamide, taxanes or trastuzumab, preexisting heart disease etc. If the female gender is found to be a risk factor for anthracycline cardiotoxicity in humans, this would be an important addition to our knowledge base. Perhaps, future investigations should focus on human studies. A unique feature of anthracycline cardiotoxicity in human is a very wide range of susceptibility to cardiotoxicity across different patients: with some patients, developing overt heart failure at relatively lower cumulative doses and some developing only minor or no apparent changes in cardiac function despite larger cumulative doses. Obviously, there are yet unidentified biochemical, proteomic or metabolomics pathways conferring susceptibility or protection from anthracycline cardiotoxicity. Clinically observed cardiotoxicity results from the summation of these factors. Future studies could focus on identifying these biomarkers. With the identification of these factors, one may be able to identify any individual's susceptibility to anthracycline cardiotoxicity prior to starting the therapy or soon after administration of first few doses of anthracyclines. The review team recommends that the NCTR collaborate with clinicians treating large number of patients with anthracycline agents for clinical indication or investigators studying these patients in clinical

research studies. We lack specific biomarkers to predict predisposition to anthracycline cardiotoxicity or to detect this at any early stage after starting therapy. Therefore, most clinical algorithms rely on serial monitoring of cardiac function by one of the several different techniques to detect an incipient deterioration of cardiac function as a harbinger of further more severe cardiac damage. Currently, this early sub-clinical deterioration of cardiac function is the stopping point for further therapy with anthracyclines. However, this approach is far from perfect, expensive and inconvenient for the patients. Development of more specific proteomic, genomic or metabolomics marker for any individual patient's susceptibility to cardiotoxicity would largely overcome these limitations.

Dr. Desai also made a presentation the development of a mouse model of doxorubicin-induced delayed onset cardiotoxicity. Doxorubicin can result in left ventricular dysfunction and heart failure, sometimes years after completion of doxorubicin treatment in patients with cancers. Detection of left ventricular dysfunction and heart failure 12 months or longer post completion of doxorubicin therapy is called late onset cardiotoxicity. Mechanistically it is unclear whether this entity represents a variation of the typical, classic cardiotoxicity and heart failure, which manifests soon after completion of doxorubicin therapy or within months of completion of doxorubicin therapy or is altogether a separate entity. However, with increasing long-term survival of cancer patients, particularly in pediatric malignancy, this area is of significant clinical interest. Dr Desai presented ideas on the development of a mouse model of delayed onset doxorubicin cardiotoxicity, by giving animals smaller cumulative doses of doxorubicin with longer follow up of up to 24 weeks post last dose. The goal here was to identify biomarkers from blood samples collected in the early part of the study, which would predict late-onset heart failure in these animals. The study design is interesting. However, this design does suffer from some methodological limitations. Since all animals in this study are genetically related to each other (littermates), one would anticipate them to behave fairly similarly or uniformly in terms of changes in left ventricular function over time. One possible variation would be to use animals derived from different strains or unrelated animals to introduce genetic variation in the model and then one may expect these animals to behave differently in terms of susceptibility to doxorubicin cardiotoxicity. The panel was posed the question whether the project will help address the knowledge gaps that has prevented the development of sensitive biomarkers capable of predicting the risk of delayed-onset cardiotoxicity. The panel thinks that this is an interesting, but somewhat challenging question. Although, there are no studies of this kind in the literature and this protocol is definitely worth pursuing, there are concerns whether this study would provide the answers investigators are seeking. Perhaps, parallel long-term clinical studies in patients undergoing doxorubicin therapy for clinical indications may need to consideration.

*Tao Han*'s presentation was on the systems biology approach to identify early biomarkers of Sunitinib-induced cardiac toxicity. A large number of TKIs are in use for various malignancies such as renal cell cancer, leukemias, lymphomas, pancreatic and neuroendocrine tumors and gastrointestinal stromal tumors. TKIs act by a variety of mechanisms by blocking the effects of VEGF on tumor cells. Several of the TKIs also cause cardiotoxicity. The mechanism of TKI cardiotoxicity is not completely understood. Dr Han and colleagues plan to use a systems biology approach to investigate sunitinib cardiotoxicity. As a first step, they intend to find out the dose range of sunitinib associated with cardiotoxicity in male and female mice. Following which, they intend to treat the animals with the cardiotoxic dose for a period ranging from 3-21

days and then perform a battery of tests including cardiac troponin, echocardiography, light and electron microscopy, matrix-assisted laser desorption ionization imaging mass spectrometry (MALDI), microarrays of the heart, liver and plasma and metabolomics and proteomic studies. This is an interesting and well-designed study and is likely to provide important insights into the mechanism of sunitinib cardiotoxicity and perhaps cardiotoxicity of other TKIs as well.The panel feels that this is probably an interesting and appropriate way to study a relatively complex project. An important point to consider is that when using genetically identical animals, they are likely to respond in a uniform manner. However, in the human context, genomic diversity may contribute to the varying susceptibility to the cardiotoxicity of TKIs. It is hard to design an animal study, which can replicate the genomic, proteomic and metabolomic diversity one would encounter in the human clinical studies. Perhaps with the availability of data from this research study, reviewers would be in a better position (in the future) to recommend additional strategies or approaches.

### Theme 3: TKI Markers and Mechanisms

Dr Mattes provided an excellent review of the cell signaling and role of TKs in cell signaling and the contribution of malfunction of TKIs in oncogenesis. The presentation also provided an overview of different small molecule TKIs in use and under clinical development, and their clinical use in various tumors. The cardiotoxicity, hepatotoxicity and pulmonary toxicity of targeted cancer therapy were described in detail, including the relative of non-specificity of TKIs for binding to the target TKs. Some of the cardiotoxicity of TKIs may be related to “off target“ inhibition of other kinases such as AMP activated PKs. The DSB is involved in several projects to address a variety of key issues in the context of organ toxicity of TKIs. These studies include prediction of species, organ, age and gender specific toxicity of various TKIs based upon kinome analysis; developing a model and index of cardiotoxicity of various TKIs using pluripotent, stem cell derived cardio myocytes, and identification of early biomarkers of cardiotoxicity of TKIs. Within the above broad group, Dr. Vijay made a presentation on the differential gene expression as a possible predictor of susceptibility to TKIs organ toxicity. He is planning to address the issue of susceptibility to organ toxicity of TKIs in rats using gene expression profiling of various organ systems. Dr. Vijay is hypothesizing that differential gene expression in different organs may help to explain the differences in susceptibility of various organ systems to toxicity of various TKIs. He is approaching this using both drug-centric approach and gene centric approaches. He has already found that the kinase expression normally varies significantly across various organs systems and predictive of the organ toxicity of various TKIs. This is an ambitious data-mining project and the panel encourages this study. These studies are likely to provide critical information about mechanistic aspects of various TKIs. Dr. Vijay enquired whether inclusion of other drug properties such as $IC_{50}$, $K_d$, kinetic parameters would add value to achieve the stated goals, and how should the severity of toxic effects be handled while categorizing the toxicity of various TKIs. Inclusion of these properties are more important when the investigators move into analyzing human data. With the rodent data, the review team feels is it is important to contextualizing the data relative to other public information with non-TKI’s. For example, when “candidate” gene expression changes are observed, it is important to identify where these genes map and under what circumstances do they change.

Dr Vijayalakshmi Varma’s presentation was on the impact of obesity of pharmacokinetics, drug action and drug toxicity. With obesity becoming a major issue on a global scale, Dr Varma raises

an important question about the specific interaction between obesity and pharmacokinetics and alterations in drug toxicity. With profound changes in multiple different metabolic pathways, inflammatory biomarkers and gene expression, it is reasonable to suspect some impact of these variables in drug pharmacokinetic as well as drug toxicity. She presented data of sunitinib pharmacokinetics in an obese patient, where the plasma drug level was significantly lower compared to the non-obese patients. Obesity was observed to negatively influence the therapeutic efficacy of doxorubicin-docetaxel neoadjuvant chemotherapy in locally advanced breast cancers in a study from Poland. Dr Varma is planning to study the interaction between obesity and cardiotoxicity of anthracyclines in human. She is also planning to study the influence of diet-induced obesity in mice on drug toxicity. There was a uniform support and enthusiasm from the committee for this project. The review team was also supportive of the presenter's query whether to explore the effectiveness of doxorubicin to reduce tumor burden in obese animals? The review team was interested to know whether there was any data comparing the different obesity models (diet induced, spontaneously occurring obesity)? If not, it may be worthwhile conducting some studies using different obesity models to tease out if it the obesity (by any pathway) or the underlying biochemical, metabolic or genetic abnormalities causing obesity will result in differences in drug toxicities and therapeutic efficacies. The review team opines that the different organ systems particularly liver should be studied for susceptibility to drug toxicities as well as drug metabolism. To the question whether there was any specific high priority drugs to be investigated for specific drug safety/toxicity issues, the review team is of the opinion that to begin with, drugs with narrow therapeutic safety or efficacy ranges may be studied, because even small variation in drug metabolism and pharmacokinetics may have larger clinical impacts in these cases. Alternatively, one may select medications handled by the pathways most likely to be affected by obesity.

## Theme 4: Alternative Drug Safety Models

There were three presentations on the theme of Alternative Drug Safety Models. 1) *Using iPSC-Derived Cardiomyocytes to inform and Predict Drug-Induced Structural Cardiotoxicity* (Drs. White and Yang). This presentation focused on the promise of iPSC-derived cardiomyocytes (iPSC-CMs) for cardiac risk testing. The research offers a possible model to support a mechanism-based biomarker discovery program which could lead to better preclinical safety evaluations. The specific research goal of this group emphasized the importance of developing an *in vitro* human stem cell derived model which would increase the mechanistic understanding of tyrosine kinase inhibitor (TKI)-induced cardiotoxicity (CT). The project may also provide insight toward identifying novel biomarkers. The presentation began by describing a brief application of electrophysiology of structural CT. The presentation then continued by offering some original data using a variety of *in vitro* toxicity endpoints for several TKIs, followed by an analysis program for risk ranking, and a summary and future direction, while demonstrating the value of combined endpoints for risk prediction of structural CT. The model is acceptable particularly since the use of iPSCs has generated significant interest in toxicity testing. The importance lies in the advantage of iPSCs over other human and mouse embryonic stem cells. In general, iPSCs avoid the controversies that have plagued the handling of human embryonic stem cells in biomedical sciences, and circumvent the irrelevance of mouse stem cells and their application to human toxicology. However, the model suffers from significant shortcomings. *In vitro* behavior of these cells necessitates precise manipulation of the growth, proliferation and differentiation properties. Additionally, the process of initially establishing iPSCs is cumbersome

and not always successful. However, there are reports in the literature of other researchers (e.g., Sun et al., 2012) using these approaches when studying genetic differences. have been studying genetic  The introduction of viral vectors into human cells adds another challenge for future customization of iPSCs. Benchmarking the specificity and sensitivity thresholds of these models could be hampered by the limitations in the number of inhibitors evaluated. However, it must be mentioned that the research group is still in its early stages of investigation.  Other comments from the review team questions the predictive ability or practicality of an *in vitro* model in general. One member doubted the applicability of any *in vitro* cardiac cell model to be able to mimic the complexity of the heart. This is significant because it should not be presumed that cultured cells have the ability to simulate structural or physiological cardiac components, despite their unique features.  However, this was not a consensus opinion of the review team members. Dr. White made it clear from the beginning of the presentation that the goal of the project was aimed at developing an *in vitro* model to elucidate mechanisms underlying CT, and to identify non-invasive biomarkers to detect structural toxicity. The review team suggests that the lab expose the future perspective of the model's potential to determining the toxicity of different classes of cardiotoxic drugs, where the iPSCs may show significant improvement in the sensitivity and specificity of CT.  The model may then be used, at least initially, as an adjunct in combination with other *in vitro* procedures, as a 'screening method' for toxicity to replace LD50 studies in animals. Once the methodology is well established, the goal of determining inter-individual susceptibility becomes realistic. One of the review team members opined that these models do have utility in developing medications and for general toxicity screening as long as the caveats mentioned above align with the decision-making and at the appropriate drug development phase. Aligning responses with rank-order potency is important in toxicity assessments. In many instances, they do have advantages to in vivo rodent or large animals as long as the limitations are clearly understood.

Dr. Amy Inselman's presentation was titled, "*Human induced pluripotent stem cells – an alternative model to investigate the effects of opioid exposure on neural development*". The presentation started with the description of the status of the opioid epidemic, the risk to special populations, especially to women during pregnancy, and the potential for developmental neurotoxicity. The technical approach appeared reasonable and the incorporation of iPSCs as the basis of the model is important, as noted above. The laboratory appears to be adequately equipped and staffed, and the lab's focus appears to be pointed multiple directions. Since the project has just initiated and is still awaiting approval, now is the opportune moment (in the early stages) for the leadership needs to determine which direction is most efficient. Depending on the preliminary results, the work may follow a path of examining the effects of opioids on neural differentiation, or it be focused to assess the sensitivity of the model to the drugs, or may choose to determine the relevance of the model to the *in vivo* and public health situation. The review team cautions that investigating differentiation of stem cells to NPCs (neural progenitor cells), neurons and glial cells, is a formidable task. ,

Dr. Noriko Nakamura's presentation was titled, "*Evaluation of in vitro male reproductive toxicology models: in vitro spermatogenesis*, was presented by Noriko Nakamura. There was a summary discussion of the background of alternative model development and the presentation quickly discussed several relevant *in vitro* male reproductive toxicology models, including models for gametogenesis, spermatogenesis, organ culture systems, and microfluidic devices.

She then continued with some preliminary results from the organ culture studies, and toxicological endpoints for both *in vivo* and *in vitro* studies. The review team feels that the objectives may need better definition. Overall, the review team notes that the different efforts have resulted in specialized laboratories with deep expertise in specific areas. However, the review team feels that although the laboratories are highly engaged, the overall directions and applications may need some fine-tuning. The DSB has the proficiency and tactical position in the federal regulatory and testing arenas. The release of the landmark report, "Toxicity Testing in the 21st Century: A Vision and a Strategy (U.S. National Academy of Sciences, 2007) has precipitated a major change in the way toxicity testing is conducted. The findings envisions decreased animal usage and increased toxicity testing efficiency by proposing fundamentally new directions in toxicology. This report can be valuable to the DSB as a road-map for the development of alternative models to animal testing methods, particularly in light of advances in understanding biological responses to chemical stressors.

## Theme 5: Food Safety Technology

There were two presentations in this theme. One centered around the RAPID-B, a universal platform for real-time detection of pathogens (Dr. Buzatu) and the other focusing on bacterial serotyping using mass spectrometry (Dr. Wilkes). The RAPID-B assay is based on detecting specific pathogens using labeled antibody or oligo probe mediated flow cytometry. The work performed to date on this technology is noteworthy with research findings being published in good journals. The presentation also alluded to the fact that the technology has been tested against bacterial and protozoan pathogens as well as prions. There was also mention that the technology has been tested in a limited scale against *Mycobacterium* sp. The SpecID MS for Rapid Triage is also another technology being developed for rapid triage of bacterial isolates. It is very obvious that a significant amount of effort has gone into the development of these technology platforms. The RAPID-B has also been awarded a US patent in 2015. The researchers should explore how their expertise in flow cytometry can enhance current genomic technologies. The CFSAN and the CDC relies on genomic detection platforms such as real-time PCR and whole genome sequencing to detect and characterize pathogens. There is no doubt that these genomic technologies cannot differentiate between live and dead cells. It may be, therefore, worthwhile to explore whether the RAPID-B is applicable as a primary screening approach prior to genomic analyses. The commercial food industries in the US rely on DNA or RNA based detection routinely. It is unsure whether this could be the reason why Dr. Buzatu' s group is having difficulty in getting "market pull" for their technology even though they have completed the level 2 validation for toxigenic *E.coli* in spinach. Similar to the RAPID-B, the SpecID MS is beset with market challenges that go beyond the science and technology. The SpecID relies on pure cultures of bacterial isolates for the serological identification and characterization. One unique aspect of the technology is the use of pattern recognition software for analyzing and interpreting the MS spectra. Serological characterization of bacterial isolates is no longer the benchmark for characterizing pathogens. Even in the case of the emerging non O157 "Big Six" strains the current standard is the use of real-time PCR detection of specific DNA sequences. Additionally, the CDC currently uses whole genome sequencing (WGS) to transform pathogen detection and determining transmission patterns. (http://www.cdc.gov/pulsenet/pathogens/wgs.html). The science within the DSB in the food safety technology is of high caliber. The review team feels that the DSB may need to develop a business plan more than a scientific strategy to be able to commercialize their technologies in the

target market. They may also have to focus on the most likely market (with the most urgent need) and perform key experiments in collaboration with the private food industry. Overall, the researchers in this thematic area have deep expertise in chemistry. To address contemporary food safety issues especially in the development of technologies that are relevant to food safety issues there needs to be very close interaction and collaboration with food microbiologists. The outstanding chemistry expertise that exists within this division can be marshalled into exploiting the technology platforms in the emerging areas of metabolomics especially with reference to the behavior of probiotic cultures in foods and the state of microbial populations in fecal transplant programs. The review team commends this group for their efforts to commercialize their technologies. However, the review team also feels that there will be synergy if the DSB group collaborates with the Division of Microbiology.

## Theme 5: Computational Modeling

There were two presentations in this thematic area. Both of them focused on the use of 3D spectral data as chemical descriptors. The first, given by Dr Beger, focused on the theory behind the approach whereas the second presentation focused on the application to hERG inhibition and mutagenicity prediction. The theory focuses on using predicted NMR chemical shifts and interatomic distances to model the relationships between chemical structures and their observed biological activity. The work was originally focused solely on using the $^{13}C$ NMR spectra of compounds and the initial work was patented in 2004. The approach was enhanced with the incorporation of atom pair distances and subsequently this enhanced methodology was patented in 2011. Further extensions of the approach to include 15N and 17O atom types was done and published in 2013. The chemical shifts observed in NMR spectra are reflective of the electronic environment of the atom and thereby capture much of the local chemical environment of an atom. The electronic environment can be an important determinant in its ability to form Van der Waal's interactions or degree of reactivity. When coupled with descriptors that describe the 3D molecular shape and hence the surface accessibility of individual atoms, this makes for a potentially thorough approach to describing how a molecule might interact with a biological system or protein. Dr Beger went on to describe how the methodology had been applied to predicting phospholipidosis and the inhibition of the human ether-a-go-go (hERG) potassium ion channel in a collaboration with NCATS and using data from assays developed by this center. He also spoke briefly about other research applications in the prediction of Ames mutagenicity although this was covered in more detail in the second presentation by Dr Slavov. The modelling performance for hERG was significantly impacted by the classification of the experimental data into a binary system of active or inactive. While this approach may be appropriate for some biological effects such as mutagenicity where the concentration where the effect is seen is considered irrelevant in a risk assessment context, most decisions in the determination of risk with drugs and chemicals in general are based on a comparison of exposure at which toxicity is observed versus the exposure administered to human subjects or patients. Therefore in most cases there is a need to consider predictions of the concentration where adverse biological effects would be anticipated. In addition, when using a classification approach that is based on a concentration effect then a significant proportion of the available information will be lost when the data is converted from a continuous scale to a binary one and thus will impact on the performance of the model in general. For example, there is a distinct difference between a molecule that inhibits the hERG channel activity at a concentration of 1nM versus a compound that only inhibits the activity of the hERG channel at a concentration of 1uM.

In using a binary classification system the subtle but important structural determinants of activity will be invisible to the modelling algorithm and hence unidentifiable. Finally, Dr Beger described methods for defining the applicability domain of the 3D-QSAR models and presented data from models for estrogen receptor binding data taken from the ToxCast data set. According to OECD guidelines, one of the requirements for a QSAR model to be considered acceptable in a regulatory setting is the ability to determine if a chemical lies within the applicability domain of the model. This is considered to be important for determining the validity of a prediction based on the chemistry domain that is represented within the training set as extrapolation of statistical correlations beyond this are inappropriate and more likely to be error prone. That said, the challenge here is how to appropriately represent the extent of the applicability domain of a model. In fact, a model that is based on an understanding of the chemical or biological mechanism(s) may well apply to chemicals that are significantly different from those represented in the training set. This presents a significant hurdle to model developers when trying to show that the algorithm or model can accurately determine when a prediction is considered to be unreliable. However, in the data presented by Dr Beger, the prediction accuracy for those chemicals deemed to be outside the applicability domain are generally less accurate than those that are within it. However, this can be test set dependent and may be misleading. More recent developments in the machine learning world have looked at statistical methods, such as the conformal prediction framework, to describe the degree of confidence in prediction which ultimately may be a more useful and pragmatic solution to the applicability domain problem.

Dr Slavov's presentation focused on the Ames mutagenicity prediction. This work was in collaboration with an external company, Lhasa Ltd. which is internationally recognized for its expertise in genotoxicity. In contrast to hERG or Estrogen receptor models which describe a chemical's binding interaction with a discrete protein, this data set focused on a chemical's ability to cause point mutations in DNA which are for the most part driven by chemical reactivity and hence represents a fundamentally different type of problem in terms of modelling. Dr Slavov discussed ways in which the data used for modeling was sub-categorized for modeling purposes but was converted back to a binary classification at the end of the process. The exact process for converting the data to a continuous variable was not adequately described. The review team cautions that the modelers be very familiar with the data that they are using and how it is generated as it is important to take the limitations of the biological data into consideration when using it. The results of the modeling exercise for Ames mutagenicity are broadly in line with those seen with commercially available systems. The fact that no system has achieved significantly better than 80% accuracy for models built using public domain data and used to predict most test sets might be more indicative of the quality and accuracy of the data used in the building of the models rather than a limitation of the modelling technique or the chemical descriptors used. The partial least squares (PLS) used in this modelling approach while still considered a reasonable statistical methodology for QSAR modelling, is not considered to be state-of-the-art. Therefore, it would be fruitful to look at other machine-learning approaches such as support vector machines or random forest. Dr Slavov presented the workflow for the modelling process and mentioned that that computational time required per compound was in the order of several minutes albeit on a relatively old machine. Although not unmanageable, this processing time would significantly impact its application to where 100's or 1000's of structures needed to be evaluated in a single analysis.

The following are answers to questions posed to the review team. 1) *Would the 30-SDAR models of hERG and PLD be useful for pharmaceutical companies and regulatory agencies?* Potentially yes, although there are two major limitations that would need overcoming. Example, for most toxicity endpoints it is important to be able to predict the concentration at which an effect is likely to be observed rather than a simple yes/no answer. Also, the computational time for processing a structure is quite long as so would limit it's utility in a screening setting where 100s or 1000s of compounds might need to be evaluated in the model.
*2) What other toxicological endpoints should we model using SDAR and 30-SDAR technologies?* Predictions of unwanted pharmacological effects, cytotoxicity or any other discrete toxicological mechanism would be beneficial.
*3) What other companies or organizations should we collaborate with?* This would very much depend on the purpose of the collaboration but the European Bioinformatics Institute has created the ChEMBL database which might be useful source for predictions.
*4) Should the two distinct data sets of primary and secondary amines be combined into one and modeled together?* This might be worth trying as the modelling technique should be able to separate the relative importance of the descriptors as they describe the electronic status
*5) Is it good idea to use an intermediary endpoint i.e. transforming the categorical variable into a continuous variable?* Only in cases where this done appropriately and with a full understanding of the data and its significance. This does not appear to be the case with the way it was described for the mutagenicity data set. However, using the continuous IC50 for hERG and then using the predicted value to subsequently classify compounds as active versus inactive based on a threshold is a valid approach.
*6) What would you consider a better model – a model that predicts better or a model that explains better (provides structural/mechanistic interpretation)?* The ideal scenario is a model that can do both. Obviously, the best model is one that is more accurate in its predictions however, experience suggests that recipients of predictions that are not modelers themselves are more comfortable with and accepting of a prediction that can be interpreted and explained or rationalized provided that it gets it right much more often than it gets it wrong. The review team recognizes the DSB's deep expertise in data analysis and computational modeling and their collaborations. The research at the DSB in computational modeling has the possibility of leading to advances in accuracy for some QSAR problems. The review team, however, recommends that the computational modeling team take additional steps to understand the biology of the data that is being modeled. Furthermore it is recommended that the researchers try to seek the simplest solution to a complex problem. This is because descriptors that are computationally slow to generate will limit their application in high volume data sets or, where a high-throughput model is required.

## Theme 6: Precision Medicine

There were two presentations on the theme of Precision Medicine. The first was a genomic assessment of sex- and age-related differences to treatment-induced adverse events in rats by Dr. Fuscoe. The second was a study on the impact of obesity on drug pharmacokinetics, action and toxicity by Dr. Varma. Both of these thematic issues are peripherally associated with human precision medicine initiatives and are currently understudied fields of research. Further, both of these areas of research align with the FDA mission of speeding innovations that make medicines more effective, because they provide information that is relevant to individual patients. Understanding the intrinsic and extrinsic variables inherent to individual patients could

ultimately improve real-world medication effectiveness. The patient variables of gender, age, and obesity are not typically part of sponsor-conducted efficacy trials, yet critical to medication tailoring. The NCTR is a logical institution to not just participate in, but lead this field of research, and the review team is highly supportive of these efforts.

The first studies presented, led by Dr. Fuscoe, began with mining PharmaPendium gene expression data to find temporal and gender-specific changes in the expression of genes known to be involved in drug metabolism and transport. The gene most differentially expressed in these data mining exercise was cyp3a2, which is the rat orthologue of CYP3A4 in humans. If the 2500-fold greater expression level in male relative to female rats manifested in proportional changes in human CYP3A4 activity, the implications would be substantial, as CYP3A4 is at least partially involved in the metabolism of roughly half of marketed medications. Therefore, answering this obvious question on preclinical species translatability raised by the data is of interest to the scientific community. Dr. Fuscoe further mined the literature to identify 41 drugs metabolized by sexually dimorphic P450 enzymes in adult rats with the proposed next step of testing toxicity endpoints in primary hepatocytes. The preliminary results with terfenadine were not consistent with 2500-fold expression difference in cyp3a2 *in vivo*, and *in vitro* cyp3a2 levels were not presented. The problem as stated is important, but the preliminary data and addition of a new variable (model cell culture system with different drug metabolizing capability) make the 'next steps' difficult to justify as proposed. If the goal is to publish these finding, the approach likely will produce publishable manuscripts relative to in *vitro/in vivo* gene expression and metabolic capability of rat hepatocytes. If the goal is personalized medicine as indicated the next steps will need to be adjusted to account for translatability.

Hepatic differences in drug metabolism are well-documented (Waxman et al., 2009) (*Waxman DJ, Holloway MG. Sex Differences in the Expression of Hepatic Drug Metabolizing Enzymes. Molecular Pharmacology. 2009;76(2):215-228*.), and in general, thought to be less prominent in humans than in rodents. For example, the gene expression trend observed with cyp3a2 may actually be reversed in humans (Scandlyn et al., 2008) (*Scandlyn MJ, Stuart EC, and Rosengren RJ (2008) Sex-specific differences in CYP450 isoforms in humans. Expert Opin Drug Metab Toxicol 4 413-424*). To move toward the FDA's mission of improving medication effectiveness, one approach would be to move toward the human clinical situation in future research rather than away from the clinic, and back into rat hepatocyte models. This approach would include mining the human literature for PK differences in the 41 drugs identified, and continue rat investigations *in vivo*, focusing on endpoints proximal to the question of drug metabolism relative to exposure rather than *in vitro* cytotoxicity. With the current next steps as proposed, the data will be useful to further prove that hepatocytes are transcriptionally different than rat livers, but will not likely drive NCTR toward precision medicine, which is centered upon real-world effectiveness.

The second precision medicine presentation was by Dr. Varma, who laid out a clear rationale on how obesity could affect drug toxicity through a variety of mechanisms, including pharmacokinetics. The project plan included developing a diet-induced mouse model of obesity. The specific aims centered upon phenotyping these mice, then using the new model to test if obesity exacerbates doxorubicin toxicity. There were well-planned experiments and consistent with Dr. Fuscoe's research outline, aligned with the FDA mission of speeding innovations that

make medicines more effective. To the best of the review team's knowledge this research focus of the impact of obesity on drug toxicity is very novel. This is a highly relevant topic and one that the DSB's activities could be pioneering. A deeper understanding of these interactions would ultimately protect public health and could actually lead to the design of relevant clinical trials. The review team does not question the relevance of the approach or underlying science. Rather, questions during discussions arose as to whether model development and Doxorubicin toxicity are the most efficient utilization of what appear to be very limited resources for this project. The review team feels that the Doxorubicin approach seems to be force fit to align with other DSB activities. The questions Dr Varma is looking to answer do not require insult with a model toxin. Multiple animal models of obesity are available, and Dr. Varma did mention several of these in her slides. Furthermore, focusing on chronically administered drugs, such as those to treat metabolic syndrome, may be a better group of agents to pursue in testing the hypothesis and producing results that would impact obese, high-risk healthcare patients. Chronic medications are not typically dosed based on body surface area, and therefore represent medication classes most likely influenced by the effects of obesity. Drugs intended to treat diabetes, cardiovascular disease, and dyslipidemia administered chronically in dose-driven paradigms might offer a good starting point, especially if exposure were the primary readout. The effects of obesity on drug toxicity requires the measurement of multifactorial endpoints, so question may be best answered using simple endpoints and multiple drugs, which is an alternative approach to the proposed Doxorubicin model. Both approaches could be accomplished in parallel. However, the issue of resource allocation may be a challenge. The review teams notes the following areas of strength in this thematic area, namely, 1) pioneering proposal to use animal models to address important precision medicine questions, 2) efficient and thorough transcriptomic data mining, and 3) the overall project scope aligns with precision medicine, yet differentiates from rest of field that focuses on genetics. The review team also noted that there was no apparent indication of expansion of precision medicine initiatives since these projects appear to staffed minimally. In addition, genomic efforts of the NCTR in the future appear to move away from, and not towards translatability and precision medicine