CAROLINA CENTER FOR COMPUTATIONAL TOXICOLOGY: Experimental and computational tools for NexGen safety assessments

Funded by U.S. EPA Cooperative Agreement STAR RD 83516601

Progress Report for Project Period 07/01/12 - 09/30/13

Report Date: 10/24/13

Ivan Rusyn, Principal Investigator

Table of Contents

Overview

The ***long-term objective*** of our ongoing research program is to advance the science and practice of toxicology by (i) filling critical gaps in our knowledge of the toxicity mechanisms, (ii) incorporating the population-based screening methods into the practice of toxicity testing, (iii) developing reliable computational models and tools that address specific existing challenges in hazard identification, and (iv) engaging with the stakeholders to increase the impact of our work.

Objective 1: Develop a quantitative high-throughput screening (qHTS) approach to probe differential chemical effects in a population-based in vitro system.

Objective 2: Provide the computational toxicology solutions for risk characterization in NexGen assessments with a focus on point-of-departure and population variability.

Objective 3: Develop cheminformatics-based, as well as enhanced chemical-biological, models of in vivo reproductive and developmental toxicity that rely on concomitant exploration of chemical descriptors and population-based screening data.

The major aims and objectives have not changed from the original application. Extensive progress has been made on all specific objectives as detailed below. In addition, Quality Assurance activities consisted of annual technical audits on the work under each specific objective. The annual meeting with the External Advisory Board was held on 30 January, 2013 and the report from advisors was received on February 4, 2013. The response to advisor's recommendations is also included in this report.

Progress Report

Specific objective 1: Develop a quantitative high-throughput screening (qHTS) approach to probe differential chemical effects in a population-based in vitro system.

In an experiment that was funded by STAR RD83382501, in collaboration with NCATS and NIEHS/NTP we successfully screened 1086 human lymphoblast cell lines, representing 9 populations from 5 continents, in a cell viability assay with 179 diverse environmental chemicals at 8 concentrations (Abdo et al. 2012; Abdo et al. 2013). The extent of inter-individual variability was less than 10 fold for about 2/3 of the 179 compounds; however, some compounds showed more than a 100-fold range. Across all 179 compounds, clustering of EC_{10} profiles revealed similarity of responses in cell lines according to the genetic ancestry. For individual chemicals, population differences were evident for about 40% of the compounds. Heritability analysis, using populations with parent-child trios, showed that cytotoxicity responses of 22 chemicals were significantly heritable (~20-40%). Because cell lines used in this study have been densely genotyped or sequenced, analyses of genome-wide association, and association pathway analyses, were performed to identify significant associations between variants/genes/pathways and cytotoxicity. Across the 179 compounds, 142 polymorphisms were identified as suggestive of the genetic association. Of the 1086 screened cell lines, basal gene expression (RNA sequencing-based) data is publicly available for 344 cell lines and we performed correlation analysis of this data with cytotoxicity across the 179 chemicals to identify additional genes and pathways that may also be associated with inter-individual differences in toxicity.

While such population-based human *in vitro* model offers exceptional opportunities for evaluating the potential hazard, mode of action, and population variability in response to chemicals, limitations remain that require further assessment to increase the utility of the information obtained from this model. Thus, a follow-up study (Abdo et al. 2014) was designed to address two of these limitations: restricted metabolic capacity of lymphoblasts and the potential to screen complex mixtures. We selected lymphoblast cell lines from 4 racially and geographically diverse populations based on availability of genome sequence and basal RNA-seq. For the mixtures experiment, 146 cell lines were exposed to pesticide mixtures (organochlorine pesticide environmental mixture extracted from a passive surface water sampling device, or a mixture of 36 currently used pesticides). For the drug/metabolite experiment, we exposed 331 cell lines to Carbamazepine, Sulfamethoxazole, or two major metabolites of each drug. Concentration-response cell viability data was used to derive a 10% effect (or no-effect) level for each compound/cell line. We found that a mixture of pesticides currently in-use was less toxic but more variable than a mixture of organochlorine pesticides. A significant correlation between pesticide mixtures indicates that cells that were sensitive to one pesticide mixture were likely sensitive to the other. Significant population differences were found for the currently used pesticide mixture. For drug/metabolite pairs, we observed a wide range of inter-individual variability in cytotoxicity. Significant pairwise correlations of EC10 were found for each drug and its two metabolites. We found statistically significant and/or suggestive associations between genetic variants and cytotoxicity of each drug/metabolite. These data provide the opportunity to establish population-based confidence intervals in cytotoxicity, as well as probe candidate susceptibility pathways.

As an alternative way of addressing the limitations of lymphoblast cells with regards to metabolism and target-specific toxicity, we have initiated collaboration with *Molecular Devices* and *Cellular Dynamics* who are world leaders in high-content/high throughput cellular imaging and induced pluripotent stem cell (iPSC)-based *in vitro* models, respectively. In these studies (Sirenko et al. 2013a; Sirenko et al. 2013b; Sirenko et al. 2013c) we used iPSC-derived hepatocytes and cardiomyocytes to screen large (100+) libraries of chemicals and determined how the multi-parametric assessment of concentration response toxicity phenotypes can be used to make hazard-based rankings.

Specific objective 2: Provide the computational toxicology solutions for risk characterization in NexGen assessments with a focus on point-of-departure and population variability.

a) Developing computational solutions for estimation of the population variability in toxicity

The *1000 Genomes Toxicity Screening* data (Abdo et al. 2013) is largest ever population-scale experiment. To utilize the power of this data and to explore the value of such data for computational toxicology solutions for hazard identification and dose-response, we partnered with Sage Bionetworks (Seattle, WA) to use this data for one of DREAM (Dialogue for Reverse Engineering Assessments & Methods) Challenges (Erickson 2013). The DREAM8 challenges were held in the period from June 10, 2013 through September 15, 2013. We provided the participants in the ***NIEHS-NCATS-UNC Toxicogenetics Challenge*** (http://dx.doi.org/10.7303/syn1761567) with access to the 1000 Genomes Toxicity Screening data which included ~200 chemicals evaluated for cell viability in lymphoblastoid cell lines (1000 Genomes Project Consortium 2010) from 1000+ individuals of 9 distinct subpopulations across Europe, Africa, Asia, and the Americas. These data were paired with the publicly

available genomic data from these cell lines, including DNA sequencing profiles and RNAseq-based transcriptomic data.

In sub-challenge 1, the participants were asked to predict inter-individual variability in *in vitro* cytotoxicity based on genomic profiles of individual cell lines. For each compound, participants were challenged to predict the absolute values and relative ranks of cytotoxicity across a set of unknown cell lines for which genomic data is available. For sub-challenge 2, the task was for each compound, predict the concentration at which median cytotoxicity would occur, as well as inter-individual variation in cytotoxicity, described by the 5-95th%ile range, across the population. Each prediction was scored based on the participant's ability to predict these two parameters within a set of compounds excluded from the training set.

The NIEHS-NCATS-UNC Toxicogenetics Challenge attracted a "crowd" of ~250 researchers who used these data to elucidate the extent to which adverse effects of compounds can be inferred from genomic and/or chemical structure data. There were 99 models submitted by 35 teams for sub-challenge 1, and 85 models by 23 teams for sub-challenge 2. Final announcement of the winners of the challenges will be made at the 2013 DREAM conference that will held on November 8-12 in Toronto, Canada in conjunction with the RECOMB/ISCB Conference on Regulatory and Systems Genomics.

b) Developing computational solutions for organ-specific toxicity using iPSCs

In our iPSC-based studies (Sirenko et al. 2013a; Sirenko et al. 2013b; Sirenko et al. 2013c), we also determined what utility these data have for risk characterization in NexGen assessments. We focused on point-of-departure data as information for hazard classification. We assessed prediction accuracy of the individual parameters, as well as the multi-parametric data that was integrated by the means of point-of-departure information into a single prediction using ToxPi software developed in collaboration with US EPA/NCCT with funding from STAR RD83382501 (Reif et al. 2013).

c) Developing computational solutions for estimation of the point-of-departure

Chemical hazard assessments necessarily vary based on data availability and the type of risk management decision they support. While much recent attention has been on the use of high-throughput toxicological data for screening and prioritization, assessments that support toxicity guidance values or standards still rely on epidemiological and *in vivo* experimental data. Such assessments, including Integrated Science Assessments and Integrated Risk Information System Toxicological Reviews, are highly data, time, and resource-intensive, and cannot be realistically expected for most environmental chemicals. In response to the need to develop default approaches to support risk estimation for chemicals lacking chemical-specific information, we are developing the Conditional Toxicity Value (CTV) Predictor in collaboration with EPA/NCEA (Wignall et al. 2013a). This project aims to address the challenge of creating an approach to support risk estimation for chemicals lacking chemical-specific information. CTV tool uses chemical properties and limited experimental data to predict toxicity values (such as the oral slope factor, inhalation unit risk, reference dose and concentration), as well as points of departure. Toxicity potency ranking can also be generated for groups of chemicals. CTV predictions rely on a new comprehensive database of existing toxicity values, the associated points of departure (with benchmark doses calculated where feasible) and other experimental data. The approach combines QSAR and regression modeling, and incorporates OECD principles for model building and external cross-validation. We believe this project has potential

to be used by various stakeholders. The EPA SAB/BOSC commented that "*the EPA's effort to develop the concept of Conditional Toxicity Values (CTV) is particularly noteworthy because it incorporates consideration of new toxicity testing methods, and offers the potential to create screening or interim risk values for large numbers of chemicals of concern*" (US EPA 2012).

To enable development of CTV for point-of-departure values (*e.g.*, BMD or LOAEL), we had to calculate BMD values for a large number of chemicals that have been evaluated in human health assessments and had dose-response datasets. We aimed to apply a standardized process for conducting BMD modeling to reduce inconsistencies in model fitting and selection and to identify study design features affecting BMD modeling fit acceptability (Wignall et al. 2013b). We evaluated dose-response data (880 datasets) for 352 environmental chemicals with existing human health assessments. We calculated benchmark doses (1 standard deviation or 10% response, BMD10/1SD) for each chemical in a standardized way with pre-specified criteria for model fit acceptance. We identified study design features associated with successful models. Our approach was successful for 255 (72%) of chemicals. Batch-calculated BMD10/1SD values were significantly and highly correlated with points of departure previously used in human health assessments, with median values similar to reported no-observed-adverse-effect-levels (NOAELs). We observed a significant trend of increasing model viability with increasing number of dose groups. At the same time, choices of dose level and spacing had little effect. We conclude that BMD10/1SD values can be calculated in a standardized way for use in health assessments on a large number of chemicals/critical effects. This facilitates exploration of health effects across multiple studies of a given chemical, or when chemicals need to be compared, providing greater transparency and efficiency than current approaches. Lessons learned from successful or failed benchmark dose-response modeling can inform study design and selection for future assessments.

<u>d) Developing computational solutions for cloud-based development of human health assessments of chemicals</u>

As one of the solutions that can be offered to the organizations tasked with developing human health assessments of chemicals, we are developing HAWC (Health Assessment Workspace Collaborative, https://hawcproject.org/), a modular, cloud-ready, informatics-based system to synthesize multiple data and information (Shapiro et al. 2013). Developed in collaboration with EPA/NCEA, this system seamlessly integrates and documents the overall workflow from literature search and review, data extraction and evidence synthesis, dose-response analysis and uncertainty characterization, to creation of customized reports. Crucial benefits of such a system include improved integrity of the data and analysis results, greater transparency, standardization of data presentation, and increased consistency. By including both a web-based workspace for assessment teams who can collaborate on the same assessment rather than share files and edits, and a complementary web-based portal for reviewers and stakeholders, all interested parties have dynamic access to completed and ongoing assessments.

Specific objective 3: Develop cheminformatics-based, as well as enhanced chemical-biological, models of in vivo reproductive and developmental toxicity that rely on concomitant exploration of chemical descriptors and population-based screening data.

Endocrine disrupting chemicals (EDCs) are a growing public concern due to their adverse effects on human and wild life. In order to develop *in silico* predictors to identify Estrogen Receptor (ER)-mediated EDCs, we collected from public databases and scientific literature a

large number of ER ligands along with reported relative binding affinity to ERα and/or ERβ (546 compounds for ERα and 137 compounds for ERβ). A novel multi-task learning (MTL) QSAR modeling approach was applied to develop models capable of predicting the binding affinity of ligands to both ER subtypes. Compared with conventional single-task learning (STL) models, MTL models significantly improved the predictive accuracy for ERβ binding affinity (R^2 increased from 0.32 to 0.53) while keeping the high predictive accuracy of ERα models. In addition, as a complementary approach, docking studies were performed on a set of ER agonists/antagonists (67 agonists/39 antagonists for ERα and 48/32 for ERβ) and corresponding presumed decoys/non-binders (2570/1448 for ERα and 1000/1000 for ERβ). These compounds were docked to four protein conformations: ERα agonist, ERα antagonist, ERβ agonist, and ERβ antagonist. Results showed that all four conformations were capable of discriminating their corresponding ligands from presumed decoys/non-binders, with ERα agonist conformation being the best in separating ERα agonists from antagonists. Virtual screening of an uterotrophic dataset validated that the consensus of MTL QSAR and docking models had the highest enrichment power. Virtual screening of the EPA Tox21 library yielded a prioritized list of 286 putative estrogenic compounds for future *in vitro* and *in vivo* tests on endocrine disruption (Zhang et al. 2013).

Similar studies on the thyroid hormone receptor beta (THRβ) are nearing completion. In this case, we have collected and curated data on compounds that have been characterized as binders or non-binders of the THRβ as well as those tested as agonists or antagonists of the THRβ. In contrast to previous QSAR models of THRβ we have considered two sets of compounds that interact with two binding sites at THRβ, *i.e.*, Ligand Binding Domain (LBD; 129 compounds) and Ligand-Dependent Activation Function-2 (AF-2) Domain (181 compounds). Both continuous and binary QSAR models were built for both LBD and AF-2 datasets. The accuracy of the continuous model for the LBD dataset was relatively low (external R^2=0.48), which prompted us to develop binary QSAR model for the same dataset; the latter achieved the external Balanced Accuracy as high as 73%. The continuous QSAR model for the AF-2 dataset had relatively high accuracy of R^2=0.7 so we found it to be unnecessary to develop a binary model for this dataset. In a separate effort, we have also conducted docking studies with a dataset including 57 agonists and 15 antagonists of LBD domain as well as with another dataset comprising 210 antagonists of the AF-2 domain of the THRβ receptor. We have demonstrated that one of the crystal structures of the THRβ receptor (PDB code 1Q4X) was capable of effectively differentiating agonists from antagonists; similar docking studies of the antagonists of AF-2 domain showed that we were unable to differentiate these molecules from a library of presumed decoys (*i.e.*, compounds expected not to bind to AF-2). Both QSAR and docking models will be employed for virtual screening of the Tox21 library (similar to ER models described above).

We also developed a new method termed Chemical Biological Read Across (Low et al. 2013) which infers each compound's toxicity from both chemical and biological analogues whose similarities are determined by the Tanimoto coefficient. Classification accuracy of CBRA was compared to that of classical RA and other methods using chemical descriptors alone or in combination with biological data. Different types of adverse effects (hepatotoxicity, hepatocarcinogenicity, mutagenicity, and acute lethality) were classified using several biological data types (gene expression profiling and cytotoxicity screening). CBRA-based hazard classification exhibited consistently high external classification accuracy and applicability to diverse chemicals. Transparency of the CBRA approach is aided by the use of radial plots that

show the relative contribution of analogous chemical and biological neighbors. Identification of both chemical and biological features that give rise to the high accuracy of CBRA-based toxicity prediction facilitates mechanistic interpretation of the models. This approach was highlighted in Chemical Watch (http://chemicalwatch.com/16098/chemical-biological-read-across-may-offer-improved-accuracy?q=chemical%20biological)/

Activities for Subsequent Reporting Period

In *Specific Objective 1*, we will:

- Finalize the analysis of the 1000 Genomes Project screening
- Finalize the analysis of the population-wide experiment with mixtures and drug/metabolite pairs
- Further explore the utility of iPSC models for population-based high-content/high-throughput screening by developing additional collaborations with *Cellular Dynamics* who are establishing iPSCs from hundreds of individuals with sequenced genomes

In *Specific Objective 2*, we will:

- Work with the winners of the NIEHS-NCATS-UNC Toxicogenetics Challenge to develop user-friendly and publicly available computational approaches based on the best-performing models
- Finish development of chemical structure- and biological data-based CTV
- Finish and deploy HAWC
- Continue working with US federal agencies and other stakeholders to improve functionalities in HAWC

In *Specific Objective 3*, we will:

- Work with EPA Office of Research and Development and Office of Chemical Safety and Pollution Prevention on applying ER models to chemical prioritization for in vivo screening
- Finish virtual screening studies using both QSAR and docking models of the THRβ receptor. Submit the manuscript that describes the development of QSAR and docking models for THRβ; initiate and complete the development of similar models for other ED-related receptors such as AR.

Changes in Key Personnel

Dr. Fred Wright has assumed a new position of Director at the Bioinformatics Research Center at North Carolina State University. He also was appointed professor of Statistics and Biological Sciences at North Carolina State University. Because Dr. Wright will retain an adjunct faculty status at UNC-Chapel Hill, no changes in his involvement in this project are expected.

Quality Assurance

A Quality Assurance Project Plan (QAPP) was drafted and finalized in January 2013. The QAPP includes sections for each specific objective and task under each objective. Responsibilities of each project member were identified. QA Manager, Dr. Karin Yeatts, conducted the annual technical audits of the projects in January 2013. The audit identified a need

for minor corrective action in all objectives and the results were distributed to the principal investigators as well as the EPA Project Officer, Dr. Pasky Pascual in January 2013. A report to the External Advisory Board was made on January 30, 2013.

References:

1000 Genomes Project Consortium. 2010. A map of human genome variation from population-scale sequencing. Nature 467(7319): 1061-1073.

Abdo N, Marlot P, Pirmohamed M, Shea D, Wright FA, Rusyn I. 2014. Utilizing human population based in vitro model to investigate pesticide mixtures and drug/metabolite pairs. In: Society of Toxicology Annual Meeting. Phoenix, AZ.

Abdo N, Xia M, Kosyk O, Huang R, Sakamuru S, Austin C, et al. 2012. The 1000 genomes toxicity screening project: Utilizing the power of human genome variation for population-scale in vitro testing. In: Society of Toxicology Annual Meeting. San Francisco, CA.

Abdo N, Xia M, Kosyk O, Huang R, Sakamuru S, Brown C, et al. 2013. The 1000 genomes toxicity screening project: Utilizing the power of human genome variation for population-scale in vitro testing. In: Society of Toxicology Annual Meeting. San Antonio, TX.

Erickson BE. 2013. Crowdsourcing toxicity prediction. Chem Eng News 91(26): 18.

Low Y, Sedykh A, Fourches D, Golbraikh A, Whelan M, Rusyn I, et al. 2013. Integrative chemical-biological read-across approach for chemical hazard classification. Chem Res Toxicol 26(8): 1199-1208.

Reif DM, Sypa M, Lock EF, Wright FA, Wilson A, Cathey T, et al. 2013. ToxPi GUI: an interactive visualization tool for transparent integration of data from diverse sources of evidence. Bioinformatics 29(3): 402-403.

Shapiro AJ, Cook N, Ross PK, Fox J, Cogliano V, Chiu WA, et al. 2013. Web-based benchmark dose modeling module as a prototype component of an informatics-based system for human health assessments of chemicals. In: Society of Toxicology Annual Meeting. San Antonio, TX.

Sirenko O, Crittenden C, Callamaras N, Hesley J, Chen YW, Funes C, et al. 2013a. Multiparameter in vitro assessment of compound effects on cardiomyocyte physiology using iPSC cells. J Biomol Screen 18(1): 39-53.

Sirenko O, Cromwell EF, Crittenden C, Wignall JA, Wright FA, Rusyn I. 2013b. Assessment of beating parameters in human induced pluripotent stem cells enables quantitative in vitro screening for cardiotoxicity. Toxicol Appl Pharmacol: in press.

Sirenko O, Hesley J, Rusyn I, Cromwell EF. 2013c. High-content assays for hepatotoxicity using induced pluripotent stem cell (iPSC)-derived cells. Assay Drug Dev Technol: in press.

US EPA. 2012. Implementation of ORD Strategic Research Plans: A Joint Report of the Science Advisory Board and ORD Board of Scientific Counselors. Washington, DC.

Wignall JA, Muratov E, Fourches D, Tropsha A, Woodruff T, Zeise L, et al. 2013a. Conditional Toxicity Value (CTV) predictor for generating toxicity values for data-sparse chemicals. In: Society of Toxicology Annual Meeting. San Antonio, TX.

Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, et al. 2013b. Standardized Benchmark Dose Calculation: Opportunities to Inform Science-Based Decisions in Human Health Assessments. Environ Health Perspect: in review.

Zhang L, Sedykh A, Tripathi A, Zhu H, Afantitis A, Mouchlis VD, et al. 2013. Identification of putative estrogen receptor-mediated endocrine disrupting chemicals using QSAR- and structure-based virtual screening approaches. Toxicol Appl Pharmacol 272(1): 67-76.

Annual Review by the External Advisory Board
February 4, 2013

Dear Dr. Rusyn:

This letter provides a summary of comments by the External Advisory Board (Drs. Motsinger-Reif, Tice, Whelan, and Daston) on the progress of the North Carolina Center for Computational Toxicology. The summary is based on our collective review of the materials that you and your colleagues prepared and the review we attended at the Univ. of North Carolina on 30 January, 2013. Overall, we were very impressed with the portfolio of projects, the excellent progress that you are making, and the level and quality of interaction with U.S. EPA. We were also pleased to have the opportunity to meet and hear from students in your program. They are to be commended on the quality of the presentations they made.

We have several general comments for your program, as well as a number of suggestions for the individual projects that were presented.

General Comments:

1. The research program does an excellent job of applying cutting-edge computation and biotechnology to real-world problems and will have positive impacts on public health. While all the projects that we heard about are progressing well and will make a positive contribution, we believe that the highest priority should be given to projects that rely on the unique computational and statistical capabilities of the Center.

   The Board believes that there is opportunity for even greater collaboration, across projects within the Center as well as externally. A key example might be the in the endocrine disrupter area, where there are two projects within the Center and numerous external activities focused on predicting toxicity based on nuclear hormone receptor interactions.

2. The projects are all making great progress, and the students who presented summaries each provided very ambitious, but not very specific lists of next steps. We urge you to provide greater specificity on these. As you do this, please keep in mind the suggestion we make in the first bullet point above.

3. We commend you on the robust quality assurance process that you have implemented. On the computational side, the Git environment that has been set up will prove useful for shared development of such ambitious software projects.

Project-Specific Comments:

Thousand Genomes Screening Project:

The Board was impressed with the initial study. It provides a great deal of information that will be useful in addressing one of the great uncertainties in human health risk assessment: the nature and magnitude of variability in response across the population. The Board makes the following suggestions for further research:

1. Research that supports the relevance of lymphocytes as surrogates for systemic responses to toxicants
2. Evaluation of chemicals that provide even greater diversity of mode-of-action, and analysis of which modes-of-action are represented in lymphocytes
3. Exploration of experimental readouts beyond ATP decrease/ cytotoxicity that may be indicative of a greater range of biological responses
4. Multivariate analysis of gene variant contributions to susceptibility (or resistance), especially a determination of whether genes with similar functions or in a common pathway are associated with differential response
5. Deeper dives on a selected number of individual chemical responses to ensure that the results are consistent with the available literature on the toxicity of those chemicals
6. Integration of other high throughput and high content data generated on the same compounds into the overall analysis of compound similarities and differences

Endocrine Disruption Hazard Modeling and Virtual Screening:

We were impressed with the extent of work that was done to characterize interactions of a large number of small molecules with two estrogen receptor isoforms. This work provides an additional tool, with ultra-high throughput, for assessing the potential of new chemicals to be estrogenic or anti-estrogenic, a tool that can be used both for screening and green chemistry design.

1. Consider developing docking and QSAR models for additional receptors (e.g. AR).
2. Consider developing models that predict the kinetics of ligand binding, especially dissociation rates
3. Consider linking the docking/QSAR results with downstream events as the first step in creating a quantitative computational model of estrogenic response.
4. Develop a better means of communicating the biological relevance of receptor docking, given that 80% or so of the tested chemicals docked with the receptor.
5. Consider using the docking/QSAR results to identify compounds that are not able to react with a specific nuclear receptor.
6. Consider additional ways to collaborate with other groups who are interested in screening for endocrine disruption, particularly groups that generate wet lab data on receptor binding.

Conditional Toxicity Value (CTV) Predictor:

The intention of this project is to provide a comprehensive computational tool for the prediction of reference values for untested chemicals, using structural analog and QSAR approaches. We agree that such a tool could be useful, and make suggestions for possible improvements.

1. Consider adding threshold of toxicological concern (TTC) as a first tier for chemicals that do not have sufficient toxicity data. The literature supporting TTC has increased over the past few years, as has regulatory guidance (especially in Europe) on its use. Further analysis of TTC data sets also provides useful information on the chemical features associated with high or low potency, which could be used in the analog algorithms
2. Consider relying on point-of-departure data rather than Reference Dose values as the basis for providing predictions of toxicity for analogs
3. Consider including additional databases that include toxicity data beyond the endpoints represented in the current version of CTV

4. Consider adopting expert-based rules that have been published and validated (e.g., Wu et al, 2010 Reg Tox Pharm 56:67 and Blackburn et al 2011 Reg Tox Pharm 60:120) as a basis for improved evaluation of analog suitability

Standardizing Effects of EDCs Listed in a DG Environment Database:

The Board felt that this small project provided an interesting case study on how an Adverse Outcome Pathway (AOP) framework can be used to organize causally-linked information at different levels of biological organization to profile a chemical's potential to act as an endocrine disrupter. It also suggested a useful start on a standardized terminology for endocrine-mediated health effects, and provided practical recommendations on improvements to adverse outcome pathway analysis. While it wasn't clear if you foresee any additional work in this area, we will be on the lookout for opportunities to apply this work to larger efforts concerning the elucidation of ED-related AOPs and on the development of mode of toxicity ontologies that are likely to commence in the near future.

HAWC: Health Assessments Workspace Collaborative:

This project provides a virtual teamspace for the development and review of risk assessments. At present it is tailored towards EPA NCEA needs. We believe that the project will be very useful for NCEA, and could be adapted for use by others who conduct risk assessments. We have no specific suggestions for this project. Concern was raised that the project may be a little too applied to be among the highest priorities for the Center, but is still within scope.

It was a pleasure to be able to review your program. It is making important contributions that will advance toxicology and risk assessment.

Sincerely,
Alison Motsinger-Reif
Ray Tice
Maurice Whelan
George Daston

Response to the Comments by the External Advisory Board

General Comments:

1. The Center PIs strive to further improve both internal and external communications. Internally, regular meetings of PIs and research staff have been instituted. Drs. Rusyn and Tropsha, and staff working on Objectives 2 and 3 meet bi-weekly. Drs. Rusyn and Wright and their students and staff working on Objectives 1 and 2 meet at least bi-weekly and frequently smaller meetings take place even more regularly. With respect to external communications, especially in the endocrine disruptors area, these concern our partnerships with various divisions at US EPA, and our collaborators at NIEHS/NTP and Sage/DREAM. Specific examples include, but are not limited to, close interactions with the National Center for Computational Toxicology (EPA/ORD) and Office of Science Coordination and Policy (EPA/OCSPP) to share our ER models and make predictions on chemical libraries in ToxCast Phase II and Tox21. In collaboration with the National Center for Environmental Assessment (EPA/ORD), we are developing CTV and performed BMD modeling on a large number of dose-response datasets. In collaboration with NCEA (EPA/ORD) and the Office of Health Assessment and Translation at the National Toxicology Program we are developing additional functionalities in HAWC tool.
2. We regret that due to the limitations in time at the meeting we did not articulate the future plans adequately. Accordingly, we have provided specific plans for each Objective in this annual report.
3. We continue to use Git and other resources to assure quality and robustness of our code.

Project-specific comments:

Thousand Genomes Screening Project:

The Board was impressed with the initial study. It provides a great deal of information that will be useful in addressing one of the great uncertainties in human health risk assessment: the nature and magnitude of variability in response across the population. Our response to the suggestions made by the Board is as follows:

1. Our follow-up experiments are designed to address the relevance of lymphocytes as surrogates for systemic responses to toxicants through the experiments with drug/metabolite pairs and two pesticide mixtures. In the next year we will be performing more in-depth analysis of the data that was collected in April-June 2013.
2. With regards to the "evaluation of chemicals that provide even greater diversity of mode-of-action, and analysis of which modes-of-action are represented in lymphocytes" we plan to explore large-scale testing of chemical mixtures. Our pilot with two mixtures is encouraging and cells were frozen to potentially explore the molecular mechanisms of toxicity, in addition to the *in silico* exploration of the MOA through GWAS and basal gene expression/toxicity correlation analysis.
3. With regards to the "exploration of experimental readouts beyond ATP decrease/cytotoxicity that may be indicative of a greater range of biological responses", we plan to perform additional molecular experiments with cells that were frozen in our mixtures/drug-metabolites experiment.

4. With respect to the "multivariate analysis of gene variant contributions to susceptibility (or resistance), especially a determination of whether genes with similar functions or in a common pathway are associated with differential response", we note that these are exactly the types of analyses we are exploring through GWAS and explorations of the correlations between basal gene expression and variability in cytotoxicity.
5. Concerning "deeper dives on a selected number of individual chemical responses to ensure that the results are consistent with the available literature on the toxicity of those chemicals," we are exploring such connections and will include such analyses in our presentations and manuscripts.
6. With regards to the "integration of other high throughput and high content data generated on the same compounds into the overall analysis of compound similarities and differences." We plan to explore what data is available from Tox21 screening. We have established an MoU with ToxCast/Tox21 and will have access to the information on chemicals and assays.

Endocrine Disruption Hazard Modeling and Virtual Screening:

The Board was very positive about the outcome of our preliminary studies of the ER receptor ligands using both QSAR and docking approaches. The Board saw particular value in employing our models for virtual screening of new chemicals to prioritize theme for the experimental investigation. The Board also made important suggestions concerning future directions of our research; our comments to these suggestions are as follows.

1. Following the Board recommendation we have developed new models of the THRβ receptor. These models can be used to predict new ligands as well as discriminate agonists versus antagonists of the receptor. Additional studies with other endocrine receptors (such as AR) are planned.
2. The problem of predicting the kinetics of ligand binding especially the dissociation rate is very challenging; it requires advanced molecular simulations approaches. We currently lack both resources and expertise to conduct such investigations.
3. So far, we only had access to data on receptor binding in vitro; however, there is recent data on compounds whose interactions with both AF-2 and LBD domains of THRβ receptor are measured by analyzing the downstream effects.
4. The question concerning the biological relevance of docking may be caused by confusion about specificity of docking results. Indeed, we have demonstrated that the success in discriminating between agonists and antagonists by docking is hard to achieve; in fact our studies of both ER and THRβ receptors suggest that not every crystallographic structure of the receptor could be used to discriminate the two ligand classes successfully. When the latter happens, the model actually effectively discriminates a small fraction of the dataset (*i.e.*, much smaller fraction than 80% of chemicals dock well)
5. Indeed, a good docking model affords highly accurate discrimination between binders and non-binders so the prediction of non-binders is implicit for our best models.
6. We have made preliminary contact with Dr. Kip Guy from the St Jude Children's Hospital who has developed an efficient HTS assay for the THRβ receptor (Johnson et al, J Biomol Screen. 2011 16(6):618-27) and who agreed to test our predictions of AF-2 binding ligands experimentally.

Conditional Toxicity Value (CTV) Predictor:

1. The suggestion for finding the potential interaction points between CTV and TTC is an excellent one. We have analyzed existing literature on TTC and will present a plan of how CTV and TTC may be mutually informative at the next external advisory board meeting.
2. The recommendation that we rely on "point-of-departure data rather than Reference Dose values as the basis for providing predictions of toxicity for analogs" is being implemented. Because few chemicals had derived BMDs and most of the point-of-departure data for chemicals that were evaluated in human health assessments is NOAEL, we derived BMD values for a large number of compounds (see Progress Report). Using these values and available NOAELs we are developing POD-based CTV models as well.
3. With regards to the recommendation that we "consider including additional databases that include toxicity data beyond the endpoints represented in the current version of CTV", we note that such work will be performed, indeed, once we finalize our current CTV efforts.
4. The Board commented that we shall "consider adopting expert-based rules that have been published and validated (e.g., Wu et al, 2010 Reg Tox Pharm 56:67 and Blackburn et al 2011 Reg Tox Pharm 60:120) as a basis for improved evaluation of analog suitability." This is a great suggestion but it will be our focus in year 3 of the project.

Standardizing Effects of EDCs Listed in a DG Environment Database:

We appreciate the encouragement and excitement that the Board had for our work in this area. However, we do not foresee any additional work in this area for the time being due to the limitations in funding. We will be, indeed, "on the lookout for opportunities to apply this work to larger efforts concerning the elucidation of ED-related AOPs and on the development of mode of toxicity ontologies that are likely to commence in the near future."

HAWC: Health Assessments Workspace Collaborative:

The Board commented that this project "is tailored towards EPA NCEA needs." While we have organized the modules and case studies in HAWC based on IRIS and PPRTV programs, we have been expanding the utility of HAWC beyond NTP. Our ongoing collaboration with NTP/OHAT on including study quality evaluation functionalities is an example of how the modules in HAWC can be used by a variety of stakeholders from the government to the industry.

Annual Report Summary

Date of report: prepared October 21, 2013 for reporting period 07/01/12 - 09/30/13.

EPA Agreement Number: #RD83516601

Title: Carolina Center for Computational Toxicology: Experimental and computational tools for NexGen safety assessments

Investigators:
Ivan Rusyn, PI
Fred Wright
Alex Tropsha
Karin Yeatts, Quality Assurance Manager

Institution: University of North Carolina at Chapel Hill

Research Category: computational toxicology

Project Period: July 1, 2012 - June 30, 2015

Objective of Research:

The objective of the Carolina Center for Computational Toxicology is to advance the science and practice of toxicology by (i) filling critical gaps in our knowledge of the toxicity mechanisms, (ii) incorporating the population-based screening methods into the practice of toxicity testing, (iii) developing reliable computational models and tools that address specific existing challenges in hazard identification, and (iv) engaging with the stakeholders to increase the impact of our work..

Progress Summary/Accomplishments:

In ***Specific objective 1*** our goal is to develop a quantitative high-throughput screening (qHTS) approach to probe differential chemical effects in a population-based in vitro system. To this end, we have been following up on the success of our collaboration with NCATS and NIEHS/NTP in which we successfully screened 1086 human lymphoblast cell lines, representing 9 populations from 5 continents, in a cell viability assay with 179 diverse environmental chemicals at 8 concentrations (Abdo et al. 2012; Abdo et al. 2013). A follow-up study (Abdo et al. 2014) was designed to address two limitations of a lymphoblast-based in vitro screening model: restricted metabolic capacity of lymphoblasts and the potential to screen complex mixtures. We screened two environmental pesticide mixtures (organochlorine pesticide environmental mixture extracted from a passive surface water sampling device, or a mixture of 36 currently used pesticides) and drug/metabolite pairs (Carbamazepine, Sulfamethoxazole, or two major metabolites of each drug). These data provide the opportunity to establish population-based confidence intervals in cytotoxicity, as well as probe candidate susceptibility pathways. In addition, as an alternative way of addressing the limitations of lymphoblast cells with regards to metabolism and target-specific toxicity, we collaborated with Molecular Devices and Cellular Dynamics, companies that are world leaders in high-content/high throughput cellular imaging

and induced pluripotent stem cell (iPSC)-based in vitro models, respectively. In these studies (Sirenko et al. 2013a; Sirenko et al. 2013b; Sirenko et al. 2013c) we used iPSC-derived hepatocytes and cardiomyocytes to screen large (100+) libraries of chemicals and determined how the multi-parametric assessment of concentration response toxicity phenotypes can be used to make hazard-based rankings.

In ***Specific objective 2*** our goal is to provide the computational toxicology solutions for risk characterization in NexGen assessments with a focus on point-of-departure and population variability. First, we are developing computational solutions for estimation of the population variability in toxicity by utilizing the power of the 1000 Genomes Toxicity Screening data (Abdo et al. 2012; Abdo et al. 2013). Specifically, we partnered with Sage Bionetworks (Seattle, WA) to use this data for one of DREAM (Dialogue for Reverse Engineering Assessments & Methods) Challenges. In sub-challenge 1, the participants were asked to predict inter-individual variability in in vitro cytotoxicity based on genomic profiles of individual cell lines. For each compound, participants were challenged to predict the absolute values and relative ranks of cytotoxicity across a set of unknown cell lines for which genomic data is available. For sub-challenge 2, the task was for each compound, predict the concentration at which median cytotoxicity would occur, as well as inter-individual variation in cytotoxicity, described by the 5-95th%ile range, across the population. Each prediction was scored based on the participant's ability to predict these two parameters within a set of compounds excluded from the training set. The NIEHS-NCATS-UNC Toxicogenetics Challenge attracted a "crowd" of ~250 researchers who used these data to elucidate the extent to which adverse effects of compounds can be inferred from genomic and/or chemical structure data. There were 99 models submitted by 35 teams for sub-challenge 1, and 85 models by 23 teams for sub-challenge 2. Final announcement of the winners of the challenges will be made at the 2013 DREAM conference that will held on November 8-12 in Toronto, Canada in conjunction with the RECOMB/ISCB Conference on Regulatory and Systems Genomics.

Second, we are developing computational solutions for organ-specific toxicity using iPSCs. We used the dose-response data from iPSC-based studies to assess prediction accuracy of the individual parameters, as well as the multi-parametric data that was integrated by means of point-of-departure information into a single prediction using ToxPi software.

Third, we are developing computational solutions for estimation of the point-of-departure. In response to the need to develop default approaches to support risk estimation for chemicals lacking chemical-specific information, we are developing the Conditional Toxicity Value (CTV) Predictor in collaboration with EPA/NCEA (Wignall et al. 2013a). CTV is an approach to support risk estimation for chemicals lacking chemical-specific information. CTV tool uses chemical properties and limited experimental data to predict toxicity values (such as the oral slope factor, inhalation unit risk, reference dose and concentration), as well as points of departure. Toxicity potency ranking can also be generated for groups of chemicals. CTV predictions rely on a new comprehensive database of existing toxicity values, the associated points of departure (with benchmark doses calculated where feasible) and other experimental data. The approach combines QSAR and regression modeling, and incorporates OECD principles for model building and external cross-validation. To enable development of CTV for point-of-departure values (e.g., BMD or LOAEL), we applied a standardized process for conducting BMD modeling to reduce inconsistencies in model fitting and selection and to identify study design features affecting BMD modeling fit acceptability (Wignall et al. 2013b). We evaluated dose-response data (880 datasets) for 352 environmental chemicals with existing

human health assessments. We calculated benchmark doses (1 standard deviation or 10% response, BMD10/1SD) for each chemical in a standardized way with pre-specified criteria for model fit acceptance.

Fourth, we are developing computational solutions for cloud-based development of human health assessments of chemicals. To this end, we are developing HAWC (Health Assessment Workspace Collaborative, https://hawcproject.org/), a modular, cloud-ready, informatics-based system to synthesize multiple data and information (Shapiro et al. 2013). HAWC seamlessly integrates and documents the overall workflow from literature search and review, data extraction and evidence synthesis, dose-response analysis and uncertainty characterization, to creation of customized reports. Crucial benefits of such a system include improved integrity of the data and analysis results, greater transparency, standardization of data presentation, and increased consistency. By including both a web-based workspace for assessment teams who can collaborate on the same assessment rather than share files and edits, and a complementary web-based portal for reviewers and stakeholders, all interested parties have dynamic access to completed and ongoing assessments.

In ***Specific objective 3*** we develop cheminformatics-based, as well as enhanced chemical-biological, models of in vivo reproductive and developmental toxicity that rely on concomitant exploration of chemical descriptors and population-based screening data. In order to develop in silico predictors to identify Estrogen Receptor (ER)-mediated endocrine dedisruption, we collected from public databases and scientific literature a large number of ER ligands along with reported relative binding affinity to ERα and/or ERβ. A novel multi-task learning QSAR modeling approach was applied to develop models capable of predicting the binding affinity of ligands to both ER subtypes. In addition, as a complementary approach, docking studies were performed on a set of ER agonists/antagonists and corresponding presumed decoys/non-binders. Virtual screening of an uterotrophic dataset validated that the consensus of MTL QSAR and docking models had the highest enrichment power. Virtual screening of the EPA Tox21 library yielded a prioritized list of 286 putative estrogenic compounds for future in vitro and in vivo tests on endocrine disruption (Zhang et al. 2013). Similar studies on the thyroid hormone receptor beta are ongoing.

In addition, we developed a new method termed Chemical Biological Read Across (Low et al. 2013) which infers each compound's toxicity from both chemical and biological analogues whose similarities are determined by the Tanimoto coefficient. CBRA-based hazard classification exhibits consistently high external classification accuracy and applicability to diverse chemicals and diverse "biological" datasets. Transparency of the CBRA approach is aided by the use of radial plots that show the relative contribution of analogous chemical and biological neighbors. Identification of both chemical and biological features that give rise to the high accuracy of CBRA-based toxicity prediction facilitates mechanistic interpretation of the models.

Future Activities:

In ***Specific Objective 1***, we will finalize the analysis of the 1000 Genomes Project screening; finalize the analysis of the population-wide experiment with mixtures and drug/metabolite pairs; and further explore the utility of iPSC models for population-based high-content/high-throughput screening by developing additional collaborations with Cellular Dynamics who are establishing iPSCs from hundreds of individuals with sequenced genomes.

In ***Specific Objective 2***, we will work with the winners of the NIEHS-NCATS-UNC Toxicogenetics Challenge to develop user-friendly and publicly available computational approaches based on the best-performing models; finish development of chemical structure- and biological data-based CTV; finish and deploy HAWC; continue working with US federal agencies and other stakeholders to improve functionalities in HAWC.

In ***Specific Objective 3***, we will work with EPA Office of Research and Development and Office of Chemical Safety and Pollution Prevention on applying ER models to chemical prioritization for in vivo screening; and finish development of QSAR and docking models for THRβ.

Publications (#)/Presentations (*)

*Abdo N, Marlot P, Pirmohamed M, Shea D, Wright FA, Rusyn I. 2014. Utilizing human population based in vitro model to investigate pesticide mixtures and drug/metabolite pairs. In: Society of Toxicology Annual Meeting. Phoenix, AZ.

*Abdo N, Xia M, Kosyk O, Huang R, Sakamuru S, Austin C, et al. 2012. The 1000 genomes toxicity screening project: Utilizing the power of human genome variation for population-scale in vitro testing. In: Society of Toxicology Annual Meeting. San Francisco, CA.

*Abdo N, Xia M, Kosyk O, Huang R, Sakamuru S, Brown C, et al. 2013. The 1000 genomes toxicity screening project: Utilizing the power of human genome variation for population-scale in vitro testing. In: Society of Toxicology Annual Meeting. San Antonio, TX.

#Low Y, Sedykh A, Fourches D, Golbraikh A, Whelan M, Rusyn I, et al. 2013. Integrative chemical-biological read-across approach for chemical hazard classification. Chem Res Toxicol 26(8): 1199-1208.

*Shapiro AJ, Cook N, Ross PK, Fox J, Cogliano V, Chiu WA, et al. 2013. Web-based benchmark dose modeling module as a prototype component of an informatics-based system for human health assessments of chemicals. In: Society of Toxicology Annual Meeting. San Antonio, TX.

#Sirenko O, Crittenden C, Callamaras N, Hesley J, Chen YW, Funes C, et al. 2013a. Multiparameter in vitro assessment of compound effects on cardiomyocyte physiology using iPSC cells. J Biomol Screen 18(1): 39-53.

#Sirenko O, Cromwell EF, Crittenden C, Wignall JA, Wright FA, Rusyn I. 2013b. Assessment of beating parameters in human induced pluripotent stem cells enables quantitative in vitro screening for cardiotoxicity. Toxicol Appl Pharmacol: in press.

#Sirenko O, Hesley J, Rusyn I, Cromwell EF. 2013c. High-content assays for hepatotoxicity using induced pluripotent stem cell (iPSC)-derived cells. Assay Drug Dev Technol: in press.

*Wignall JA, Muratov E, Fourches D, Tropsha A, Woodruff T, Zeise L, et al. 2013a. Conditional Toxicity Value (CTV) predictor for generating toxicity values for data-sparse chemicals. In: Society of Toxicology Annual Meeting. San Antonio, TX.

[#]Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, et al. 2013b. Standardized Benchmark Dose Calculation: Opportunities to Inform Science-Based Decisions in Human Health Assessments. Environ Health Perspect: in review.

[#]Zhang L, Sedykh A, Tripathi A, Zhu H, Afantitis A, Mouchlis VD, et al. 2013. Identification of putative estrogen receptor-mediated endocrine disrupting chemicals using QSAR- and structure-based virtual screening approaches. Toxicol Appl Pharmacol 272(1): 67-76.

Supplemental Keywords:
Bioinformatics, biostatistics, computational toxicology, QSAR, ToxCast, high throughput screening