BIOMED 21

A Human Pathways-Based Approach to Disease Research

NIH Conference Center
Bethesda, 26-27 June 2017

WORKSHOP *Flash* REPORT

NIH NIEHS

HUMAN TOXICOLOGY PROJECT CONSORTIUM

THE HUMANE SOCIETY OF THE UNITED STATES

HUMANE SOCIETY INTERNATIONAL

U.S. federal agencies, biomedical scientists and stakeholders examined ongoing efforts to elucidate the molecular underpinnings of human disease and discussed opportunities for enhanced coordination and prioritization of human-relevant approaches with the goal of improved translation of research results from the bench to the clinic.

Despite increasing sophistication in technology as well as in our understanding of biology, the issue of failures in clinical translation of biomedical research still looms large. This is tangibly evident, as the number of new drugs have halved every 9 years since 1950[1], despite investment of billions of dollars into these efforts and the promise of the new technologies. The global community has been aware of the shortcomings of current approaches to disease models and drug development for some time. To this end, the BioMed21 initiative and global scientific workshop series[2,3,4,5] was launched by Humane Society International to bring together all stakeholders, with the overall aim of understanding how to leverage existing projects to develop human systems biology-based roadmaps for health research. The U.S. workshop was hosted by the National Institutes of Health (NIH) with organizational support and backing from the National Institute of Environmental Health Sciences, the Human Toxicology Project Consortium, and The Humane Society of the United States.

Session 1: Setting the stage

Keynote lecture: ***Dr Chris Austin****, NIH/NCATS*
Speakers: ***Dr Francois Pognan****, Novartis;* ***Dr Warren Casey****, NIH/NIEHS;* ***Dr Bruce Cuthbert****, NIH/NIMH;* ***Dr Suzanne Fitzpatrick****, FDA;* ***Dr Rebecca Clewell*** *ScitoVation;* ***Dr Ian Adcock****, Imperial College London;* ***Dr Daniel Levner****, Emulate Inc.*

Standardization of data reporting, in line with current Food and Drug Administration (FDA) requirements for clinical trial data, could enable analysis of the dormant

[1] Scannell et al, 2012, Nat Rev Drug Discovery Diagnosing the decline in pharmaceutical R&D efficiency. 11:191 200. doi:10.1038/nrd3681
[2] Langley G, et al. Environ Health Perspect. 2015;11:A268-72.
[3] Langley G, et al. Drug Discov Today. 2016;22:327-39.
[4] https://ntp.niehs.nih.gov/go/biomed21
[5] http://www.biomed21.org

but valuable data currently in existence but often in different, incompatible formats. The advantage of organizing data to enable ontological classification is clear: with initiatives such as the European Innovative Medicines Initiative (IMI) eTOX project[6], where integration of data mining and *in silico* methods have been used to identify new likely associations between new chemical structures and toxicity, as well as helping establishing target-toxicity relationships. The organization of data can also be applied to disease, where basic mechanisms often align poorly with diagnostic criteria and here again, ontological classification would be advantageous. Using disease mechanisms to classify patients could enable more successful clinical trials and be leveraged to develop targeted therapies.[7] The development of organ-on-chip (OoC) models shows great promise for understanding normal human physiology and disease mechanisms. Human disease models of ulcerative colitis and asthma demonstrate hallmarks of the human disease that someday may lend themselves to predictions of drug efficacy. Importantly, the liver-on-a-chip developed with dog hepatocytes recapitulates aspects of the whole organ response observed in dog *in vivo* studies, providing increasing confidence that human OoC may someday be used to determine mechanisms of toxicity. This approach, possibly coupled with a tiered testing strategy that includes computational toxicology and fit-for-purpose cell based assays, may help enable *in vitro*-based safety assessment that could aid in streamlining compound development. The importance of partnerships was identified in this session, and the early engagement of the regulators will be important in developing new tools which support regulatory and research needs for the appropriate context of use. Establishing partnerships to engineer increased confidence forms one of the missions of the roadmap proposed by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and has been used successfully in the joint NIH-FDA-Defense Advanced Research Projects Agency OoC program,[8] where early discussions with regulators were necessary to identify the gaps prior to the research.

Session 2: Big data; from information to knowledge

Speakers: ***Dr Ajay Pillai****, NIH/LINCS;* ***Dr Christine Colvis****, NIH/NCATS;* ***Dr Christopher Chute****, Johns Hopkins University.*

Data accessibility was the major key in this session. The pilot project Biomedical Data Translator aims to generate a community effort to develop a highly collaborative approach to integrate multiple types of existing clinical and biomedical research data sources, including disease symptoms, drug effects, and biological data relevant to understanding pathophysiology.[9] The ultimate aim of the Translator project is to create an automated system capable of feeding disparate knowledge sources into a centralized workspace via a controller, providing novel ways to analyze existing data and enable problem solving. The LINCS project[10] is a cloud-based collection of data encompassing cell type data, perturbation data and assay data (including -omics, microenvironment and imaging studies). LINCS has adopted the Findable, Accessible, Interoperable and Re-usable (FAIR)[11] principles in order to ensure that the data are usable and this approach has also been adopted by other programs, in the U.S.[12] and further afield.[13] In addition to accessibility, data integration, linking mechanistic and phenomenological disease classifications will be required before effective data exploration can take place. However, this could lead to new classifications that take into account the impact of pathway interactions on disease severity and inform specific, targeted interventions, as is the case in Fanconi's anemia.[14]

Session 3: Existing tools to support pathway-based decisions

Participants: ***Dr Anton Simeonov****, NIH/NCATS;* ***Dr Lucie Low****, NIH/NCATS;* ***Dr Ellen Berg****, DiscoverX Corporation;* ***Dr Shannon Hughes****, NIH/NCI;* ***Dr Catherine Willett****, HSUS-HSI;* ***Dr Jeff Sutherland****, Indiana Biosciences Research Institute;* ***Dr Darrell Abernethy****, FDA*

This session addressed the utility of a team approach, using the Tox21 collaboration between FDA, NIH and

[6] http://www.etoxproject.eu
[7] https://clinicaltrials.gov/ct2/show/NCT03166501
[8] https://stemcellres.biomedcentral.com/articles/10.1186/scrt361
[9] https://ncats.nih.gov/translator
[10] http://www.lincsproject.org/
[11] https://www.force11.org/group/fairgroup/fairprinciples
[12] Wilkinson MD et al. Sci Data. 2016 Mar 15;3:160018.
[13] IMI call 12 http://www.imi.europa.eu/sites/default/files/uploads/documents/Future_Topics/IndicativeTopic_Fairification.pdf
[14] Christopher Chute- Translator and Fanconi Anemia; https://htpconsortium.files.wordpress.com/2017/03/2-3-chute_biomed21-transmed-26june2017.pdf

the Environmental Protection Agency (EPA) as an example, which has the long-term goal of developing models predictive for human toxicity. Partnerships were also seen as key for the NIH OoC program, along with adoption of this technology to the wider bioscience community. The Microphysiological Systems Consortium sees coordinated effort between funders, regulators, industry and academia, leading to the development of around fifteen different human organ systems. Using common building blocks for the chips and establishing several tissue chip testing centers to monitor the reproducibility of these systems in different labs are important features to enable confidence in the data and eventual validation of the technology. Phenotypic assays offer a complement to target-based assays, they incorporate more biology, are proximal to the clinical outcome and can capture multiple pathway activities – with a trade-off of increased cost and difficulty in validation, although less than OoC. Phenotypic assays also allow the generation of toxicity signatures, which, when associated with known adverse events in a reference database, can be used to infer mechanism of action, the pathway(s) and possible targets. Data availability is key here as the inclusion of more data (for example, results from failed clinical trials) in the reference database would enhance the utility of phenotypic profiling. Improved understanding of the complexities of cancer through increased levels of data integration is evident in the Cancer Systems Biology Consortium – a community of experimental biologists and computer modelers integrating data from clinical samples, single cell analysis, network inference, image analysis, machine learning and evolutionary theory to build, test and validate hypotheses and ideas. A systems biology approach revealed a post-translational mechanism of drug resistance[15] that was not apparent from transcriptomic analysis. Adverse outcome pathways (AOP) have been widely adopted for toxicity testing and lend themselves to the study of physiology, pathophysiology and drug discovery, facilitating molecular understanding through a framework relating biological data. Capture and storage of key events and key event relationships in the publicly available AOP Wiki[16] permits the evolution of AOP networks that indicate pathways commonality, between key events and linkages, for distinct initiating events, and may reveal novel drug targets or suggest possible avenues for repurposing existing compounds. Toxicogenomics, and specifically the analysis of gene expression signatures, offers a route to understanding the mechanism of action in unexpected toxicity reactions. These studies also emphasize the importance of data, in terms of both the quality and the relevance, and suggest that studies employing rat liver-on-a-chip as a model for rat liver/whole rat would help to increase confidence in the likely predictive ability of human OoC data. For predictive studies of drug safety and efficacy, a systems pharmacology approach would require a database(s) of integrated clinical (patient observations, genomic profile) and non-clinical (compounds and mechanistic information) data. This would also need a communication protocol between layers of information, multiple modeling tools that analyze data from different aspects (e.g. compound structure, interaction network) and an accommodating, ontology-based framework. The development of a standardized ontological infrastructure would help to bridge disciplines and organisations to bring about the integrated science that will be necessary to achieve the goal of adverse event prediction.

Session 4: Discussion & recommendations

Chair: ***Troy Seidle****, HSI*
Roundtable discussants: ***Dr Chris Austin****, NIH/NCATS;* ***Dr Frank Weichold****, FDA;* ***Dr Francois Pognan****, Novartis;* ***Dr Daniel Levner****, Emulate*

The workshop concluded with a panel discussion incorporating representatives of the major stakeholder groups (funding agencies, regulators, academia, and industry), followed by a group discussion of the critical factors both preventing and necessary for optimal collaboration and success.

Data availability and integration

Better access to data is essential, but a question is how to ensure data access and sharing? Incentives are needed, including for publication of data sets – ideally this should be entire data sets which include *all* the results. Incentives might include requirements as part of funding or publication agreements. Several ideas were explored, including that a mandate for data sharing may need to "come from the top" – an approach that could replicate the success of the Human Genome Project, including open access required by funders, dataset annotation and standardization. It was felt that the precedent of

[15] Miller et al. 2016. Cancer Discovery. Reduced Proteolytic Shedding of Receptor Tyrosine Kinases Is a Post-Translational Mechanism of Kinase Inhibitor Resistance. 6(4):382-99

[16] https://aopwiki.org

funding bodies and journals requiring that all transcriptomics data be loaded into public databases as a condition of funding/publication ideally should be extended to big data more broadly.

To enable data integration, it was felt that development of a comprehensive ontology or an overarching set of inter-operable ontologies is necessary. This would form an important initial step and could use reporting standards such as those developed for eTOX and the Monarch Initiative.

There was a recommendation to explore an international collaboration for a publicly supported public database (like the AOP Wiki), although some thought that it was preferable to have a set of databases specialized for different types of data and make these databases inter-relatable.

Data quality

A large amount of *in vivo* and *in vitro* data exist, but the data are siloed in different databases, some of which are proprietary, and in different formats, and some of which are not machine-readable. In addition, much if not most of the publically available historical data is of questionable quality. Therefore, it was generally felt that large-scale gathering of historical data might only be fruitful with a significant curation effort and focus on more recent data, and that resources may be better directed toward generating new data using better designed protocols, standardized templates, and common or inter-operable ontologies. The adoption of electronic formats such as SEND (CDISC Standard for Exchange of Nonclinical Data) will substantially reduce technical hurdles. A key challenge remains access to data submitted to regulatory agencies as part of New Drug Applications, for existing mechanisms under the Freedom of Information Act are ill-suited to large scale data curation activities.

The issue of data quality as an overarching priority came under discussion, given that data quality depends on what is defined as an acceptable level of uncertainty. Data from animal models assumes a level of confidence due to internal controls, whereas patient data has more uncertainty. A true shift in paradigm will require greater emphasis to be placed on human relevance, from top-down funding decisions to data generation to building of databases/knowledge management tools. In the meantime, some participants felt that it may be possible to carry out further exploration and publication of case studies using animal cells on chips. Comparison of these data to *in vivo* data are needed to show predictivity of the OoC and improve confidence in new technologies

Collaboration

There was general agreement that a collaborative response is required – this should be very broad and include the pharmaceutical industry and patient groups, in addition to regulators. Patient advocacy groups may be engaged to encourage open access of patient data. This also requires communication with patient groups, it was felt that the patient groups could be a powerful ally and could be better informed in order to garner their support. It is also important to broaden stakeholder participation to make toxicology an early part of R&D, not an add-on or late stage investigation. This should also include funding agency involvement to incentivize team work and open science – following the European IMI pattern for group work with various partners.

Major recommendations

- International and inter-agency collaboration is critical: formal collaboration between major organizational and funding bodies should be established
- Funding should be prioritized for researching human-based biology and promoting open access data
- Human data should be collected in collaborative, open-access high-quality databases
- Common reporting formats and common ontologies should be established for collecting and collating human biology information, from different 'omics technologies to human clinical data
- There is a need to establish formal processes for cross-sector communication
- There is an immediate need for the creation of case studies to demonstrate applications and benefits of predictive, mechanism-based approaches

Participants

Participants (approximately 150 in person and via webex), including US government (6 NIH institutes, 5 FDA centers, EPA), industry, academia, NGOs and consulting firms.

For more information, visit *biomed21.org*