BIOMED 21

A HUMAN PATHWAY-BASED APPROACH TO DISEASE AND MEDICINE

26 – 27 June 2017 | **NIH Fishers Lane Conference Center, Bethesda, MD**

White Paper

"We have moved away from studying human disease in humans… The problem is that it hasn't worked, and it's time we stopped dancing around the problem… We need to refocus and adopt new methodologies for use in humans to understand disease biology in humans." [1]

— Elias Zerhouni, MD
Former Director
U.S. National Institutes of Health

Organizing Team

Chris Austin
National Centre for Advancing Translational Sciences, National Institutes of Health
austinc@mail.nih.gov

Warren Casey
National Institute of Environmental Health Sciences, National Institutes of Health
warren.casey@nih.gov

Suzy Fitzpatrick
Center for Food Safety and Applied Nutrition, Food and Drug Administration
Suzanne.Fitzpatrick@fda.hhs.gov

Robert Kavlock
Office of Research and Development, Environmental Protection Agency
Kavlock.Robert@epa.gov

Lindsay Marshall
Humane Society of the United States, Humane Society International
lmarshall@hsi,org

Ajay Pillai
National Human Genome Research Institute, National Institutes of Health
pillaiajay@mail.nih.gov

Troy Seidle
Humane Society International
tseidle@hsi.org

Anton Simeonov
National Centre for Advancing Translational Sciences, National Institutes of Health
asimeono@mail.nih.gov

Dobrila Rudnicki
National Centre for Advancing Translational Sciences, National Institutes of Health
dobrila.rudnicki@nih.gov

Catherine Willett
Humane Society of the United States, Humane Society International
kwillett@humanesociety.org

Speakers who also contributed to this paper: Darrell Abernathy, Ian Adcock, Ellen Berg, Rebecca Clewell, Shannon Hughes, Danny Levner, Lucie Low, Francois Pognan, Jeff Sutherland

Summary

The objective of this workshop is to explore existing systems biology projects and approaches and how these projects might be better coordinated to optimally improve disease understanding and interventions.

Despite investment of billions of dollars over the past few decades, development of new drugs and other potential disease interventions remain elusive and immensely expensive. The average pre-approval cost of research and development for a successful drug is estimated to be US$2.6 billion [2] and the number of new drugs approved per billion US dollars spent has halved roughly every 9 years since 1950, decreasing around 80-fold in inflation-adjusted terms [3]. More than 90% of compounds entering clinical trials fail to gain regulatory approval, mainly as a result of insufficient efficacy and/or unacceptable toxicity, because of the limited predictive value of preclinical studies [4]. There is a growing recognition that, to increase the success rate, a stronger focus on human-relevant data is needed [1, 5].

In fact, the realization of similar limitations of existing methods for evaluating chemical safety has led the National Institutes of Health, the US Environmental Protection Agency, the US Department of Defense (Defense Advanced Research Projects Agency), as well as the European Commission and European industry, to invest hundreds of millions of dollars to develop more relevant, efficient methods to understand chemical toxicity - based on understanding and functionalizing human biological pathways [6-8]

Fundamental to these new approaches is the concept of using knowledge of biological pathways or networks (systems biology) to improve our understanding of toxicity and disease. Data-driven pathways approaches infer pathways from molecular associations, whilst curated pathway frameworks are assembled and assessed manually from data sources including published literature. The curated pathway framework concept is being developed and implemented in toxicology as "Adverse Outcome Pathways" (AOPs). A collaboration between the European Commission, the US Environmental Protection Agency and the Organization for Economic Coordination and Development (OECD) is developing the necessary software and databases to generate an "AOP Knowledgebase" [9]. The goal of the AOP Wiki project is to create a highly-curated knowledgebase of interlinking networks of biological information relating to toxicological outcomes – a systems biology knowledgebase for predicting adverse effects caused by chemical exposure [10]. These are the same biological networks that are involved in disease and are affected by drugs.

Significant investment is also being made in systems biology approaches to medicine and disease [11-16]. These approaches are focused on mining existing literature, building associative biological networks from 'omics data (e.g. genomics, proteomics, metabolomics) and pre-clinical and clinical data and in using these data to develop computational and mathematical models. These projects are making great strides in developing data mining and bioinformatics capability to collect and organize data and in creating experimental approaches and technologies that facilitate systems biology research. Both fields, chemical safety toxicology and drug discovery, which are in fact highly interrelated, could benefit greatly from coordinating resources and linking or combining data- and knowledge-bases such as the AOP Wiki.

This workshop is intended to bring representatives from several of these projects to a single venue to identify barriers and opportunities and make recommendations regarding what is needed to achieve the goal of fully implementing a human systems-biology platform for understanding disease and improving interventions.

1. Setting the stage: what is needed and why?

The 21st century has seen many pivotal advances in science and technology. Together, they offer the possibility of gaining a dynamic, systems-level and human-specific understanding of the causes and pathophysiologies of diseases, including those conditions that represent previously intractable public health challenges to our society [17]. This mechanistic understanding is a vital need, in view of ongoing translational difficulties in many areas of health research and drug discovery [3, 18-20].

Potential societal benefits that could arise from a refined, mechanistic understanding of human disease include an abbreviated critical path to new, and more effective, therapeutic interventions. Drug discovery is costly, time-consuming and complex, taking around 15 years to move from concept to product and costing over 2 billion USD [21]. It is apparent that changes are long overdue as the traditional process appears to be stagnating, in terms of investment and outputs. The total R&D expenditures for the top 10 pharma companies has not changed since 2011 and 2012. Recent estimates suggest that the overall likelihood of approval is around 9% [22]. With respect to the FDA, only 22 drugs were approved in 2016, the lowest number since 2010 [23]. Drugs fail in clinical trials for a number of reasons, but a great majority of failures are a result of insufficient efficacy and/or unacceptable toxicity, because of the limited predictive value of preclinical studies [4]. In addition, the industry halts development of many clinical candidates every year, due to toxicology findings in animals. It is unknown how many of these would in fact have been safe and efficacious in humans. Emerging technologies that allow the study of rodent and human tissues under physiological conditions may allow some of these discarded molecules to proceed under alternate development paths.

To date, preclinical discovery and development of new drugs have relied heavily on models in animals, mainly rodents, which may recapitulate only selected aspects of human physiology or disease [1,2]. The failure of animal studies to predict drug efficacy and toxicity in humans has several causes, including experimental design flaws and bias, but species variations are also significant [6]. The concordance of animal to human toxicity also varies greatly per organ, ranging from slightly less than 40% for skin, up to 90% for hematopoietic toxicities, and about 50% for hepatotoxicity [24]. Equally, preclinical toxicity prediction for human adverse events is also greatly dependent on the species and study duration [24]. There is a growing

recognition that, to increase the success rate, a stronger focus on human-relevant data is needed [7,8].

Common human diseases such as cancer, diabetes, autoimmune conditions, neurodegenerative, respiratory and cardiovascular diseases are caused by a complicated interplay of multiple genetic and environmental factors [25]. Diseases may be envisaged as the combined outcome of environment, microenvironment, phenotype, genotype, time, and other external and internal influences [26] interacting at multiple levels. Application of a pathways-based approach might allow dissection of these factors and thereby more accurate prediction of the effects of intervention within the pathway.

Identification of network(s) linking extrinsic and intrinsic factors to an adverse outcome provides a potential strategy for targeted therapeutic intervention and disease prevention. For example, a mechanistic understanding of the complex interactions between infectious pathogens and environmental toxins in the complex pathophysiology of liver cancer have informed vaccination strategies [27].

Recent studies from US (SARP, severe asthma research programme) and Europe (U-BIOPRED, Unbiased biomarkers for the prediction of respiratory disease outcomes) examining severe asthma illustrate the promise of this approach. These groups have used a combination of deep clinical phenotyping along with multi-omic analysis to define subsets of patients with asthma. SARP has focussed on transcriptomics and DNA methylation whereas U-BIOPRED has used transcriptomics, proteomics, lipidomics and metabolomics. These groups have highlighted the importance of type 2 asthma [28] as well as the that of subsets of non-type 2 asthma including inflammasome-associated asthma, metabolic dysfunction and IL-6-driven asthma [29-32]. Further analysis of these datasets will derive increased subgroups of asthmatic patients with specific disease-driving mechanisms which are not discernible from clinical features alone [33].

Systems biology approaches in cancer research have led to significant insight with regards to the heterogeneity of the tumor microenvironment and how it contributes to disease progression and drug resistance. For example, statistical modeling of unbiased proteomic data led to the identification of a new type of cell-cell communication, termed ‘reciprocal signaling’, between stromal fibroblasts and tumor cells that offers insight into drug resistance in pancreatic cancer [34]. Another recent analysis discovered drug regimens that exploit dynamic tumor sensitivities to treat acute lymphoblastic leukemia through quantification of

clonal evolution and prediction of cell fate using stochastic mathematical modeling [35]. Finally, systems biology-based pathway analysis approaches effectively identified master regulators of prostate cancer malignancy through cross-species computational analysis that effectively integrated experimental findings from mouse models and human cancer [36].

This workshop will consider the question of whether a general framework that links molecular initiating events in disease pathways and networks with adverse outcomes would provide a more predictive and effective rubric for understanding disease pathophysiology, and for targeting and evaluating new interventions.

Questions for consideration; Breakout Group Discussions Session 1

The global community has been aware of the shortcomings of current approaches to disease models and drug development for some time, including a lack of understanding of human biology ("normal" and diseased), a heavy reliance on animal studies that modestly translate to human biology, heavy expense and extended timelines. What is needed to solve the problem?

- What data are currently available and who can access these data?
- What is the role and impact of precompetitive data sharing ?
- What additional types of information and processes are needed for acquiring human data in the future?
- What are the major barriers to the pursuit of a human biology-based approach in health research, e.g., funding opportunities, journal or reviewer conservatism/bias, etc.?

2. Big Data: turning information into knowledge and knowledge into action

Much of biomedical research has a seen a recent shift away from individuals studying single end-points to multi-disciplinary teams gathering simultaneous multi-parameter data on gene, protein, and metabolite expression, cell behaviors, and organ and organism phenotypes. Multiscale data are beginning to be integrated and interpreted via systems biology tools [37]. For example, the NIH-funded eMERGE network extracts data from electronic medical records to compare whole genome scans with clinical phenotypes. It is clear that the future success and wider implementation and application of metadata from open access sources will require two basic principles: standardized data representation and straightforward techniques or software for data element mapping, along with more rigorous quality control and curation of the data.

The first phase of the trans-NIH initiative Big Data to Knowledge [14] spanned 2014-2017 and saw investment of around 200 million USD to address major data science challenges and stimulate data-driven discovery. The second phase from 2018 to 2021 focuses on access and aims to examine whether multiple datasets and computational tools can be made available for remote access by numerous individuals, utilizing the FAIR priniciples [38].

One of the recent successes of the BD2K initiative was nextstrain.org - the winner of the 2017 OpenScience Prize. This prototypic system ingests viral genome sequence data and produces phylogentic trees for display on an interactive public website. The major advantage of the program is the accelerated epidemic tracking that this approach allows. Traditionally, sequencing pathogen genomes, analyzing the data and publishing the results could take years, may require access to several disparate databases of genomic sequences and was unlikely to happen until after the epidemic had passed. The value of nextstrain.org is the almost realtime availability of these data, in a publicly accessible format.

The Library of Integrated Network-based Cellular Signatures project [7] intends to build a new, network-based understanding of human health and disease via a catalogue of so-called cellular signatures. These are the patterns of changing gene expression, proteins and other cellular processes that occur in different cells and tissues following a genetic, chemical, or environmental change. The program hypothesis is that definition of such cellular responses will lead to understanding of their mechanisms and thereby to cellular-level disease causality.

The LINCS datasets provide multiple assay results for cell lines and human primary cells following treatment with ‘perturbagens’ such as growth factors, cytokines and genetic

alterations. Interrogation of the impact of a perturbagen on multiple cell types, at multiple time points and for multiple doses is possible. The curation of a huge dataset of gene expression data in the LINCS Centre for Transcriptomics includes a search facility that allows for comparison of experimental gene expression signatures of relevance to disease with the collated data from thousands of perturbagens.

A recent success story from the LINCS project examined the role of synovial fibroblasts in rheumatoid arthritis (RA), using primary human cells, stimulated with various perturbagens and measuring protein biomarkers released [39]. The study generated three datasets of increasing complexity, with the third and final dataset made up of 50,000 data points. They assessed the effects of ten perturbagens (cytokines, growth factors and toll-like receptor ligands), in the presence and absence of ten different kinase inhibitors and using healthy and RA-derived fibroblasts, to show that the efficacy of kinase inhibition in reducing the pro-inflammatory phenotype (i.e. disrupting the cellular signature) was related to the activating ligand and not the disease status.

The study emphasizes the importance of context-dependent drug effects and the need to understand signaling networks in order to guide further drug discovery. The experimental and analytical framework for perturbation profiling is likely to have value in other disease states and whilst this is not a high throughput approach, it could allow optimization and validation of lead compounds – accelerating drug discovery.

Comprehensive integration of diverse data sets from disease names and clinical symptoms to cell types and pathways to genes, mutations, and drugs would allow the identification of novel relationships among diseases and identification of new intervention opportunities. This is the vision of the Biomedical Data Translator program (Translator), launched by NCATS in 2016 and currently in its pilot phase [40].

Questions for consideration; Breakout Group Discussions Session 2

There are several large-scale initiatives underway to mine existing data from the literature, create ontologies for curation and retrieval of this information, and to use this information to improve predictive modeling. At the same time, there is increasing awareness that much of the data is of questionable quality or relevance.

- Is the right information being captured? And how is quality of data captured or assessed?
- How best to link the output from big data projects to human information from large-scale sequencing, 'omics projects and other large data sources?
- How do we integrate new data types, such as single-cell sequencing and/or imaging, with existing data at different scales?
- Who are the users and how can the data be most effectively presented for use?
- What are we missing?

3: Current tools to support pathway-based approaches: what have we learned so far, and what existing projects/tools/information can we build from?

New technologies allow fuller, more unbiased data gathering for the creation of pathway based approaches. High-throughput screening using combinations of dozens or hundreds of cell-based assays, advanced human-specific cell and tissue based models [41] and sophisticated data assessment tools could advance the correlation of individual genetic variants with gene expression patterns, disease pathways, and phenotypic outcomes [42]. Human induced pluripotent stem cell technology offers unique access to develop cellular models of healthy, patient and disease-specific states [43]. Models derived from human stem cells have enhanced research into autism spectrum disorders [44], cardiovascular disease [45], Alzheimer's disease [46], and many other illnesses. In some instances, insights about molecular disease mechanisms and drug effects have emerged from human stem cell systems that were previously missed in nonhuman models [44]. There are, however, significant limitations that are still to be addressed, including: that most *in vitro* systems do not account for metabolism or other *in vivo* functions, and even collections of *in vitro* assays cannot simulate *in vivo* conditions; cell lines may be difficult to access and may not be as advertised; limited sourcing of primary cells and tissues; the fact that culture systems derived from iPSCs do not fully represent adult biology or disease manifestations; and the cost of developing and implementing these new systems. For pathway based approaches to be implemented, it may not be necessary to fully recapitulate *in vivo* conditions for all situations, and some of these issues are surmountable by improved technology or the application of sufficient resources, while some are basic limitations of biology, but all will need to be addressed at some level.

Large-scale, high-throughput *in vitro* assays: Tox21

Toxicology in the 21st Century [5] is a federal collaboration among EPA, NIH, including National Center for Advancing Translational Sciences and the National Toxicology Program at the National Institute of Environmental Health Sciences, and the Food and Drug Administration [47]. Using a high-throughput robotic screening system housed at NCATS and diverse assay platforms accessed by EPA and NTP, researchers are testing thousands of environmental chemicals in hundreds of cell-based assays for their potential to disrupt biological pathways that may result in toxicity. To date, the high-throughput screening effort has produced almost a hundred million experimental data points: all of these data have been made freely available to the public. Screening results give a profile of potential biological

activity of each chemical and relative activities can be used to prioritize chemicals for further in-depth investigation. Screening results can also be used to inform biological pathways involved in different types of toxicities. As examples of its utility, ToxCast data have been used to effectively screen chemicals for estrogenic activity [48], and have been combined with high-throughput modeling to prioritize chemicals for risk assessment [49].

Reconstructed tissues: organs-on-chips

Human organ-on-a-chip [50-52] culture systems provide the opportunity to develop the complexity that is currently absent from standard cell culture models. These microengineered, so-called microphysiological systems (MPS) contain cultures of human cells with dynamic fluid flow that recreates nutritional delivery, mechanical forces, and biomolecular gradients such as those experienced *in vivo*, and aim to emulate biological processes at a microscopic scale. Recent advances have seen the addition of mechanical coupling to recreate blood flow, the formation of bile canaliculi, and dynamic airflow, for models of the brain, liver and lungs, respectively [53-55]. The organ-on-a-chip approach addresses many of the shortfalls of 2D, and even some of the issues with 3D, models [56]. However, the initial start-up costs of setting up microengineered systems, the variety of device designs and culture parameters that may affect reproducibility between labs, and, thus far, the lack of validation of any devices or methodology represent future challenges [57].

The potential of the organ-on-a-chip approach, over and above 'macroscopic' cell culture, is apparent and is growing. Lung chip technology developed by the Wyss Institute at Harvard University as part of the Human-on-a-Chip project funded by the Defense Advanced Research Projects Agency (DARPA), FDA, and NIH combines design, engineering, and biology to allow the Chips to recreate the 3D living microenvironment required to obtain organ-level function. The lung-chip can replicate the function of human alveolar tissue that includes functioning immune cells, response to inflammation, and blood clot formation. The system can be used to provide lung-like tissue from healthy or diseased humans and other species for chemical testing [58]. As part of the requirements and measures of success of the DARPA project, the organs-on-chips technology platform developed by the Wyss Institite at Harvard, has been spun-out into a commercial entitity, Emulate, Inc. Emulate combines design, engineering, and biology to continue the development of new Organ-Chips systems, disease models, as well as the instrumentation and software that support the platform in collaboration with industry, acemia, and government agencies. Emulate and the Food and Drug Administration have

recently entered a Cooperative Research and Development Agreement (CRADA) to evaluate and qualify the use of Emulate's Organs-on-Chips technology as a platform for toxicology testing to meet regulatory evaluation criteria for products – including foods, dietary supplements, and cosmetics.

Mapping Biological Complexity to better understand disease

Well-characterized chemicals or drugs can be used to identify cellular response "fingerprints" that identify mechanisms of action [59] [60]. The BioMAP approach taken by DiscoverX combines signatures from panels of primary human cell-based assay to create phenotypic assay models that can be used to screen drugs for potential MoAs for both efficacy and safety. This can be applied to prioritize lead or hit compounds identified from phenotypic drug screens, selecting for the required dose and activity, and discarding cytotoxic compounds. Importantly, these phenotypic models were shown to be effective for investigating secondary activities and could be used to automatically classify toxicity mechanisms. Traditionally, target deconvolution, or molecular target identification, employs techniques, such as biochemical target-based assays or affinity-binding chromatography, that are decoupled from the biological mechanism of action and fail to differentiate off-target binding from specific activity. In contrast, these human-based phenotypic models, by enabling compound categorisation in specific mechanism classes, permit identification of off-target toxicity mechanisms and enable effective deconvolution of targets derived from phenotypic drug discovery.

In terms of complexity, cancer presents a truly heterogeneous challenge. Understanding the dynamic, multifactorial nature of this condition requires thinking and methodology from outside the traditional biological tools, needs multidisciplinary collaborations and collection, analysis and interpretation of shared data sets. A recent workshop held in April 2016 [61] brought together established and emerging researchers to showcase the potential impact of a systems biology and network mapping approach in understanding and controlling cancer. There is a pressing need for better translation of pre-clinical data to the clinic for cancer in particular, as the majority of targeted agents for oncotherapy are ineffective [62] despite promising activity in *in vitro*, *in vivo*, or pre-clinical trials [63]. The evolution of sub-clonal tumor cell populations during disease progression in cancer patients has considerable impact on treatment efficacy but is currently poorly understood. However, an –omics-based mapping of various breast cancer cell lines against primary tumor tissue has enabled pairing of cell line to primary tissue, with potential implications for the identification of tumor vulnerability [61],

although this approach does not take into account the tumor microenvironment. However, this may be addressed with recent advances in single-cell analysis. An in-depth immune profiling of renal cell carcinoma used cytometry to reveal distinct populations of immune cells that may be useful in categorizing or characterizing the tumor microenvironment [64]. Single cell RNA sequencing of cell populations extracted from melanoma has indicated impressive heterogeneity, including inter- and intra-cellular, genomic and functional - suggestive of distinct micro-environments - and revealed the presence of a drug-resistant population [65]. Systems biology approaches are vital in understanding the incredible breadth of data that arises as measurements move towards the single cell level. There is also a significant open question regarding how useful data at that depth will be for clinical use and drug development.

Mapping Pathways: Adverse Outcome Pathways as one approach

There are a number of different ways to approach the mapping of biological pathways. Pathways are associative, with some derived empirically through large-scale data generation and mapping (e.g. gene ontology networks). Others are literally drawn from historical knowledge (e.g. those in the KEGG database). The 'adverse outcome pathway' (AOP)-based concept (Figure 1) evolved from the concept of mode-of-action as it relates to chemical toxicity [66].

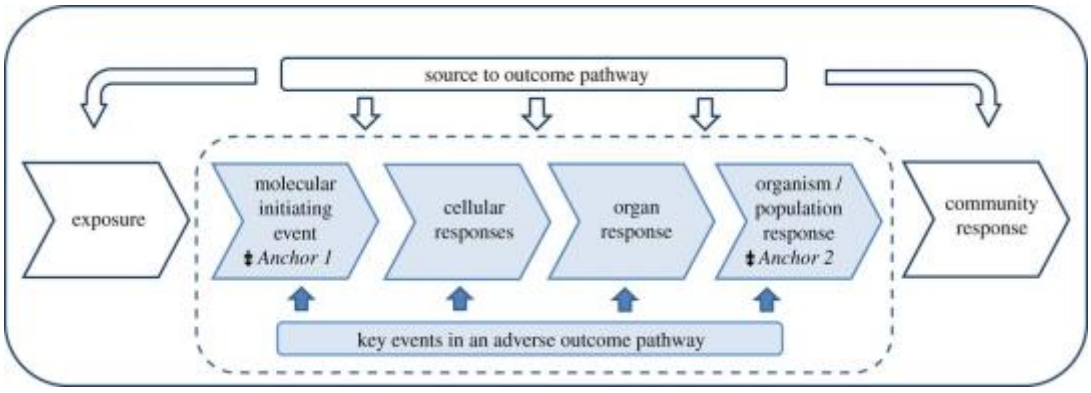

Figure 1. Simplified representation of key features of an AOP, linking two anchors – the 'molecular initiating event' (MIE) and the 'adverse outcome' at the level of an individual or population – together with intermediate 'key events'. Image reprinted from [67].

An AOP provides a framework for gathering, organizing, and evaluating biological data related to the sequence of causal relationships between a molecular initiating event (e.g., a chemical

binding a cellular receptor) and adverse outcomes connected by key events at the molecular, cellular, organ, organism, and population levels. The Organization for Economic Coordination and Development (OECD) has developed guidance for creating and evaluating AOPs [68]. OECD is also coordinating international efforts that are developing databases and knowledge-bases (curated, relational databases) to collect and document this information (the AO Wiki: https://aopwiki.org/) and that provide the basis for predictive modeling (Effectopedia: https://www.effectopedia.org/).

AOPs can be used to support weight-of-evidence evaluations of data or to design efficient testing strategies to obtain information needed for regulatory decisions [69]. OECD is developing a framework and guidance for the use of AOPs to support integrated approaches to testing and assessment (IATA) for regulatory application [70]. AOPs can also be used to improve predictive modeling for chemical safety assessment [71].

AOPs are more than associative; they provide the opportunity to substantiate causal relationships and represent the current state of knowledge for any pathway at any given time. Inclusion of a key event in the pathway indicates that it is required to achieve the phenotypic response. However, initiation of one or more of the key events may not result in the organism-level effect. Thus, an AOP defines necessity, but not sufficiency, of the key events in the progression to an adverse outcome. This supports predictive modeling that is based on the best available understanding at a given time, and, as the biological network and quantitative information continue to accumulate, the AOPs can support decisions and predictive models that have increasing greater predictive accuracy with less uncertainty. The downsides of this approach are the labor-intensive nature of the approach, that the focus is currently toxicology and regulatory applications, and that ontologies are not yet well defined to facilitate data-mining. A benefit of this framework is that it is a way for all available biological knowledge to be used to improve predictive modeling, and this could be true for any perturbation, whether it be by industrial chemicals (where it's currently being used), drugs, or disease. Another major advantage of developing AOP frameworks, is that these frameworks in themselves describe the key assay endpoints that may be used to test for chemical effect. While AOPs were developed originally to assist in evaluation of chemical safety in normal conditions, their biological underpinning supports their utility in identifying potential therapeutic targets as well.

Associative Network Mapping

Network representations have recently made the leap from social sciences, where they have been applied for decades, to systems biology. Network mappings can provide a picture of the interaction between molecules, represent the relative abundance of those molecules, and provide a molecular insight into the organization of signaling pathways, protein-protein interactions, or metabolism that would not be possible from studying individual proteins or genes. An important finding in many network analyses is that associative networks can elucidate relationships that cannot be seen when comparing single or small sets of genes, proteins or other components.

Network analysis can be applied to address how to distinguish subtle differences in mode of action [59] and for the de novo prediction of genome-wide targets. DeMAND (Detecting Mechanism of Action by Network Dysregulation), is a hybrid computational and experimental approach. This provides a network-based methodology for determining MoA and for predicting the MoA of unknown compounds and identified known and novel genome targets of vincristine and mitomycin C, and further experimentally validated these targets using human lymphoma cells [72].

Network analysis has huge potential for drug discovery and has been used to identify possible opportunities for drug repurposing [73]. For example, a network map was constructed using 3,665 FDA approved or investigational drugs, connected to their binding targets via activity profiles and then the network was interrogated using chemical structures of known drugs, to generate a score for structural similarity. This approach identified several novel relationships, one of which was later confirmed to have the predicted activity *in vivo* [74]. There is great potential to exploit this approach for complex diseases characterized by dysfunction of multiple pathways – drug network interrogation could reveal potential targets that are not obviously linked at the pathway level, enabling repurposing with reduced development time and accelerating the route to clinic. For example, a drug that has passed safety testing but failed efficacy may be repositioned via network mapping and is likely to reach the patient in a more timely and cost-effective manner. The limitations of this approach include the labour intensive nature of the work, there may be some restriction towards identification of high-affinity binding targets, and often many hypotheses are generated and only a few are tested experimentally. In some cases, other approaches (high throughput profiling in target-based assays) may have produced equivalent results.

There is inherent uncertainty in translating between pre-clinical models and humans for drug safety testing, but recent advances in -omics-base associative network modeling provides a mechanism to reduce this uncertainty. Sutherland *et al.* (2017) employed open-source databases of changes in rat liver gene expression in response to drugs to derive co-expression networks of gene sets with correlated expression patterns ('modules') [75]. Module scores for gene-sets, indicative of induction or repression of expression, were mapped to thirty-six toxicity phenotypes, derived from standard pathological assessment of liver injury – i.e. histology and biochemistry assays - and showed that pathogenic mechanisms of liver injury were associated with changes in gene expression. Importantly, analysis of the relationships between toxicity phenotype and module scores for single drug doses within the first 24 hours showed that module-phenotype associations actually preceded liver injury. This approach, when compared to *in vivo* data from rat bile duct ligation, was shown to have more discriminatory power, identifying six-fold more genes than the animal models, indicating the potential to map phenotype-gene interactions that precede liver injury and suggesting that, with the incorporation of gene-set analysis, 1-day safety studies could be employed to predict adverse outcomes. The application of rat data has the potential to reduce animal use, and may also aid in predicting human outcomes; rat-derived coexpression networks were recently shown to have predictive capacity for the identification of diagnostic gene markers for human liver disease [75].

Quantitative Systems Pharmacology (QSP) uses mathematical computer models to characterize biological systems, disease processes and drug pharmacology and facilitate greater understanding of pharmacodynamics and pharmacokinetics [76]. QSP can be used to generate biological/pharmacological hypotheses *in silico* to aid in the design of non-clinical and clinical research in order to yield more meaningful data. QSP models form useful tools for integration of prior knowledge for target identification. These are particularly important in the transition from pre-clinical studies to first-in-humans, in the move from healthy individual to the relevant patient group and in shifting from an adult to a pediatric patient population. An advance in the utility of QSP models, and evidence for their likely expansion beyond basic research, occurred in the FDA's review of the recombinant parathyroid hormone dosing regime for hypoparathyroidism, which used an open-source version of QSP models to evaluate alternative dosing strategies, the first time that QSP has been used in the regulatory arena [77]. To progress such applications, QSP models for therapeutic areas and targets that are relatively more advanced are being evaluated to understand how they might be

characterized and validated for regulatory purposes. One area that may be evolved to that point is insulin-glucose homeostasis in relationship to therapeutic targets for type 2 diabetes mellitus [78, 79]

Questions for consideration: Breakout group discussions session 3

Omics approaches have been generating an unlimited array of human data that can be leveraged to better understand biology and disease. Likewise, advanced tissue culture combined with materials science is allowing the construction of human-like tissues. "Knowledge-bases" are being created to combine all information to better describe biological understanding and support decisions as well as predictive models. How can these advanced technologies be best leveraged to improving human health?

- What is the best method for incorporating human-based -omics data into a collective knowledge-base for improved understanding of disease and better predictions for drug safety and efficacy?
- What are the pros and cons of combining medical data with the toxicological data in a platform like the OECD AOP Wiki?
- How do you integrate clinical systems biology/disease pathway knowledge into novel predictive modelling platforms?
- How can we broaden stakeholder participation (esp. to basic biological researchers)?
- What are the three main challenges for implementing systems biology understanding as a tool find practical solutions for human disease?

If there is time:

- How will we evaluate the utility of microfluidic-based systems for predicting human safety? Current approaches are anectodal and publications do not compare to simpler approaches [e.g. calculated logP and daily dose as predictors of DILI (Chen, M., Borlak, J. and Tong, W. (2013), High lipophilicity and high daily dose of oral medications are associated with significant risk for drug-induced liver injury. Hepatology, 58: 388–396. doi:10.1002/hep.26208)]

4: Coordination and support: how to make this work.

In 2015, a similar workshop discussed what might be needed to achieve fully implement mechanistic understanding of human biology to improve disease understanding and drug discovery [80]. Those workshop participants agreed that an increased reliance on human disease pathway understanding is a promising way to improve clinical efficacy of new interventions and that an adapted AOP-like approach would have great value in biomedical research [80]. Selected recommendations are listed in Figure 3.

Selected recommendations

- A 21st-century roadmap for biomedical research and/or a roadmap for each disease area need to be designed, to provide a strategic approach for utilising human-specific models and innovative technologies.
- One or two research roadmaps could be developed as prototypes, using existing knowledge and making full use of human-specific models.
- The sharing of ideas, expertise, and concepts between various research communities and clinicians needs ongoing facilitation. 21st-century tools often draw on multiple disciplines and there is frequently a disconnect between research and clinical data, which could be bridged by means of disease pathways.
- Systems biology approaches, including new bioinformatics and mathematical tools, need further support. They will help integrate multiscale data, generate better clinical disease classifications, enable disease-associated pathways to be investigated, and suggest new research directions.
- New databases are needed to integrate and share information.
- These developments require dedicated funding and policy support. Funding agreements could specify that relevant research findings are to be input into a common global knowledge base, such as the AOP Knowledge Base, and other shared, open-source platforms.
- Effective strategies are required to collect human biological material and clinical information from large patient cohorts and healthy individuals, to increase understanding of human diseases and assist the validation of new human-specific models *in vitro* and *in silico*.
- Effective data mining will be improved by harmonised standards of patient anonymisation and data protection, because this has significant implications for health research.
- Overarching strategic frameworks are essential to direct policy initiatives and funding programmes to essential areas that need further development and to coordinate related activities. These frameworks would ideally be coordinated among the EU, USA, and other key innovation economies, through a process of dialogue among all stakeholders.
- To advance 21st-century human-specific scientific progress, funding should be focussed strategically on acquiring critical human information and on developing and validating the necessary new tools, rather than on further developing animal models.

Figure 3: Excerpted recommendations from Langley et al. 2016 [80].

Questions for consideration: Group Discussions session 4:

Adequate funding and coordination of projects are critical for success of any large-scale endeavor. In this case:

- At what level(s) should this coordination occur? And who should be the "organizing body"?
- What can be done to redirect research efforts toward human biology-based approaches?
- What would incentivize industry to contribute data to populate AOP/pathway knowledge bases, etc.? How do we promote collaboration in the private sector and between the private sector and government?
- In light of emerging technologies and conceptual thinking, should there be an overarching strategic review of health research and funding frameworks and roadmap for incorporating new approaches most effectively?

References

1. Zerhouni, E.A., *Turning the Titanic*. Sci. Transl. Med, 2014. **6**: p. 221ed2.
2. DiMasi, J.A., H.G. Grabowski, and R.W. Hansen, *Innovation in the pharmaceutical industry: New estimates of R&D costs.* J Health Econ, 2016. **47**: p. 20-33.
3. Scannell, J.W., Blanckley, A.B., Boldon H., & Warrington B., *Diagnosing the decline in pharmaceutical R&D efficiency*. Nature Reviews Drug Discovery, 2012. **11**: p. 191-200.
4. Plenge, R.M., E.M. Scolnick, and D. Altshuler, *Validating therapeutic targets through human genetics.* Nat Rev Drug Discov, 2013. **12**(8): p. 581-94.
5. National Research Council, C.o.T.T.a.A.o.E.A., *Toxicity Testing in the 21st Century: A Vision and a Strategy*. 2007, Washington DC: The National Academies Press. 113.
6. SEURAT-1, *SEURAT-1 Annual Report*, in *The Proof of Concept Case Studies*, T. Gocht, and Schwarz. M., Editor. 2014: France.
7. http://www.lincsproject.org/. *The LINCS Consortium*. 04/05/2017].
8. http://www.eu-toxrisk.eu/. *EU ToxRisk*.
9. http://aopkb.org/. *AOP Knowledge base*.
10. https://aopwiki.org/. *AOP wiki*.
11. http://www.imi.europa.eu/. *Innovative Medicines Initiative*.
12. http://www.vph-institute.org/what-is-vph-institute.html. *The Virtual Physiological Human*.
13. https://ncats.nih.gov/. *National Centre for Advancing Translational Sciences (NCATS)*.
14. https://datascience.nih.gov/bd2k. *Big Data to Knowldege (BD2K)*.
15. NIH. *Cancer Systems Biology Consortium*. n.d. June 20 2017]; Available from: https://www.cancer.gov/about-nci/organization/dcb/research-programs/csbc.
16. Office of Cancer Genmoics. *CTD2: Cancer Target Discovery and Development*. 2017 June 20 2017]; Available from: https://ocg.cancer.gov/programs/ctd2.
17. van de Stolpe, A. and R.H. Kauffmann, *Innovative human-specific investigational approaches to autoimmune disease.* RSC Adv., 2015. **5**(24): p. 18451-18463.
18. Cummings, J.L., Morstorf, T., and Zhong, K.. *Alzheimer's disease drug-development pipeline- few candidates, frequent failures.* Alzheimer's Research & Therapy, 2014. **6**(4).
19. Kaitin KI, a.D.J., *Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000-2009.* Clin Pharmacol Ther. , 2011. **89**(2): p. 183-188.
20. Collins, F.S., *Reengineering translational science- The time is right.* . Science Translational Medicine, 2011. **3**(90).
21. Hughes, J.P., Rees, S. S., Kalindjian, S. B. & Philpott, K. L. , *Principles of early drug discovery.* . Br. J. Pharmacol., 2011. **162**: p. 1239-1249.
22. Thomas, D.W., Burns, J., Audette, J., Carroll, A., Dow-Hygelund, C., and Hay, M., *Clinical Development Success Rates 2006-2015*. 2016.

23. Fierce Biotech. *New drug approvals took a nosedive in 2016. Here are the 22 that made the grade.* 2017 [cited 2017 May 8]; Available from: http://www.fiercebiotech.com/special-report/new-drug-approvals-took-a-nosedive-2016-here-are-22-made-grade.
24. Greaves, P., A. Williams, and M. Eve, *First dose of potential new medicines to humans: How animals help.* . Nat Rev Drug Discov 2004. **3**: p. 226-236.
25. Lango, H. and M.N. Weedon, *What will whole genome searches for susceptibility genes for common complex disease offer to clinical practice?* J Intern Med, 2008. **263**(1): p. 16-27.
26. Gohlke, J.M., et al., *Genetic and environmental pathways to complex diseases.* BMC Syst Biol, 2009. **3**: p. 46.
27. Kirk, G.D., E. Bah, and R. Montesano, *Molecular epidemiology of human liver cancer: insights into etiology, pathogenesis and prevention from The Gambia, West Africa.* Carcinogenesis, 2006. **27**(10): p. 2070-82.
28. Woodruff, P., et al., *Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids.* Proc Natl Acad Sci U S A., 2007. **104**(40): p. 15858-63.
29. Peters, M., et al., *Plasma interleukin-6 concentrations, metabolic dysfunction, and asthma severity: a cross-sectional analysis of two cohorts.* Lancet Respir Med. , 2016. **4**(7): p. 574-584.
30. Kuo, C., et al., *A Transcriptome-driven Analysis of Epithelial Brushings and Bronchial Biopsies to Define Asthma Phenotypes in U-BIOPRED.* Am J Respir Crit Care Med., 2017. **195**(4): p. 443-455.
31. Kuo, C., et al., *T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED.* Eur Respir J. , 2017. **49**(2).
32. Rossios, C., et al., *Sputum transcriptomics reveal upregulation of IL-1 receptor family members in patients with severe asthma.* J Allergy Clin Immunol. , 2017. **17**: p. 30764-9.
33. Lefaudeux, D., et al., *U-BIOPRED clinical adult asthma clusters linked to a subset of sputum omics.* J Allergy Clin Immunol., 2017. **139**(6): p. 1797-1807.
34. Tape, C., et al., *Oncogenic KRAS Regulates Tumor Cell Signaling via Stromal Reciprocation.* Cell, 2016. **165**(4): p. 910-920.
35. Zhao, B., et al., *Exploiting Temporal Collateral Sensitivity in Tumor Clonal Evolution.* Cell, 2016. **165**(1): p. 234-246.
36. Aytes, A., et al., *Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy.* Cancer Cell, 2014. **25**(5): p. 638-651.
37. van der Sijde, M.R., A. Ng, and J. Fu, *Systems genetics: From GWAS to disease pathways.* Biochim Biophys Acta, 2014. **1842**(10): p. 1903-1909.
38. Force 11. *GUIDING PRINCIPLES FOR FINDABLE, ACCESSIBLE, INTEROPERABLE AND RE-USABLE DATA* 2017.
39. Jones, D.S., et al., *Profiling drugs for rheumatoid arthritis that inhibit synovial fibroblast activation.* Nat Chem Biol, 2017. **13**(1): p. 38-45.
40. National Center for Advancing Translational Sciences. *NCATS Funds Feasibility Assessment for Biomedical Data Translator*. 2017.

41. Singh, N., et al., *Resveratrol as an adjunct therapy in cyclophosphamide-treated MCF-7 cells and breast tumor explants.* Cancer Sci, 2011. **102**(5): p. 1059-67.
42. Muotri, A.R., *The Human Model: Changing Focus on Autism Research.* Biological Psychiatry, 2015. **79**(8): p. 642-649.
43. Bellin, M., et al., *Induced pluripotent stem cells: the new patient?* Nat Rev Mol Cell Biol, 2012. **13**(11): p. 713-26.
44. Marchetto, M.C., et al., *A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells.* Cell, 2010. **143**(4): p. 527-39.
45. Zanella, F., R.C. Lyon, and F. Sheikh, *Modeling heart disease in a dish: from somatic cells to disease-relevant cardiomyocytes.* Trends Cardiovasc Med, 2014. **24**(1): p. 32-44.
46. Choi SH, K.Y., Hebisch M, Sliwinski C, Lee S, D'Avanzo C, Chen H, Hooli B, Asselin C, Muffat J, Klee JB, Zhang C, Wainger BJ, Peitz M, Kovacs DM, Woolf CJ, Wagner SL, Tanzi RE, Kim DY., *A three-dimensional human neural cell culture model of Alzheimer's disease.* Nature, 2014. **515**(7526): p. 274-278.
47. https://ncats.nih.gov/tox21. *Toxicology in the 21st Century*. 2017 [cited 2017 May 8].
48. Browne, P., et al., *Screening Chemicals for Estrogen Receptor Bioactivity Using a Computational Model.* Environ Sci Technol, 2015. **49**(14): p. 8804-14.
49. Wetmore BA, W.J., Allen B, Ferguson SS4, Sochaski MA, Setzer RW, Houck KA, Strope CL, Cantwell K, Judson RS, LeCluyse E, Clewell HJ, Thomas RS, Andersen ME., *Incorporating High-Throughput Exposure Predictions with Dosimetry-Adjusted In Vitro Bioactivity to Inform Chemical Toxicity Testing.* Toxicol Sci, 2015. **148**(1): p. 121-136.
50. Low, L. and D. Tagle, *Organs-on-chips: Progress, challenges, and future directions.* Exp Biol Med (Maywood), 2017.
51. Low, L. and D. Tagle, *Microphysiological Systems ("Organs-on-Chips") for Drug Efficacy and Toxicity Testing.* Clin Transl Sci. , 2017.
52. Cupedo, T., A. Stroock, and M. Coles, *Application of tissue engineering to the immune system: development of artificial lymph nodes.* Front Immunol, 2012. **3**(343).
53. Booth, R. and H. Kim, *Characterization of a microfluidic in vitro model of the blood-brain barrier (muBBB).* Lab Chip, 2012. **12**(10): p. 1784-92.
54. Jang, M., Neuzil, P., Volk, T., Manz, A., and Kleber, A., *On-chip three-dimensional cell culture in phaseguides improves hepatocyte functions in vitro.* Biomicrofluidics, 2015. **9**(3).
55. Huh, D.D., *A human breathing lung-on-a-chip.* Ann Am Thorac Soc, 2015. **12 Suppl 1**: p. S42-4.
56. Ingber, S.N.B.a.D.E., *Microfluidic organs-on-chips.* NatureBiotechnology, 2014. **32**(8): p. 760-772.
57. Halldorsson, S., et al., *Advantages and challenges of microfluidic cell culture in polydimethylsiloxane devices.* Biosens Bioelectron., 2015. **63**: p. 218-231.
58. Kambez H. Benam, et al., *Small airway-on-a-chip enables analysis of human lung inflammation and drug responses in vitro.* Nature Methods, 2016. **13**(2): p. 151-157.

59. Remy, K.E., et al., *Raxibacumab augments hemodynamic support and improves outcomes during shock with B. anthracis edema toxin alone or together with lethal toxin in canines.* Intensive Care Med Exp, 2015. **3**(1): p. 9.
60. Berg, E.L., J. Yang, and M.A. Polokoff, *Building predictive models for mechanism-of-action classification from phenotypic assay data sets.* J Biomol Screen, 2013. **18**(10): p. 1260-9.
61. Archer, T.C., et al., *Systems Approaches to Cancer Biology.* Cancer Res, 2016. **76**(23): p. 6774-6777.
62. Arrowsmith, J., *Trial watch: phase III and submission failures: 2007-2010.* Nat Rev Drug Discov, 2011. **10**(2): p. 87.
63. Gan, H.K., et al., *Assumptions of expected benefits in randomized phase III trials evaluating systemic treatments for cancer.* J Natl Cancer Inst, 2012. **104**(8): p. 590-8.
64. Chevrier, S., et al., *An Immune Atlas of Clear Cell Renal Cell Carcinoma.* Cell, 2017. **169**(4): p. 736-749.
65. Tirosh, I., et al., *Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq.* Science, 2016. **352**(6282): p. 189-196.
66. Willett, C., et al., *Pathway-based toxicity: history, current approaches and liver fibrosis and steatosis as prototypes.* Altex, 2014. **31**(4): p. 407-21.
67. Madden, J.C., V. Rogiers, and M. Vinken, *Application of in silico and in vitro methods in the development of adverse outcome pathway constructs in wildlife.* Philos Trans R Soc Lond B Biol Sci, 2014. **369**(1656).
68. OECD, *Users' Handbook supplement to the Guidance Document for developing and assessing Adverse Outcome Pathways.*, in *OECD Series on Adverse Outcome Pathways*. 2016, OECD: Paris.
69. OECD, *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins.* 2014, OECD Publishing: Paris.
70. OECD, *Guidance Document on the Use of Adverse Outcome Pathways in developing Integrated Approaches to Testing and Assessment (IATA).* 2016. **Series on Testing & Assessment No. 260**.
71. Wittwehr, C., et al., *How Adverse Outcome Pathways Can Aid the Development and Use of Computational Prediction Models for Regulatory Toxicology.* Toxicol Sci, 2017. **155**(2): p. 326-336.
72. Woo, J., et al., *Elucidating Compound Mechanism of Action by Network Perturbation Analysis.* Cell, 2015. **162**(2): p. 441-451.
73. Hasan, S., et al., *Network analysis has diverse roles in drug discovery.* Drug Discov Today, 2012. **17**(15-16): p. 869-74.
74. Keiser M.J., S.V., Irwin J.J., Laggner C., Abbas A.I., Hufeisen S.J., Jensen N.H., Kuijer M.B., Matos R.C., Tran T.B., Whaley R., Glennon R.A., Hert J., Thomas K.L., Edwards D.D., Shoichet B.K., Roth B.L., *Predicting new molecular targets for known drugs.* Nature, 2009. **462**(7270): p. 175-181.
75. Sutherland, J.J., et al., *Toxicogenomic module associations with pathogenesis: a network-based approach to understanding drug toxicity.* Pharmacogenomics J, 2017: p. 1-14.

76. Leil, T.A. and R. Bertz, *Quantitative Systems Pharmacology can reduce attrition and improve productivity in pharmaceutical research and development.* Front Pharmacol, 2014. **5**: p. 247.
77. MC Peterson, M.C.a.R., M.M., *FDA Advisory Meeting Clinical Pharmacology Review Utilizes a Quantitative Systems Pharmacology (QSP) Model: A Watershed Moment?* CPT Pharmacometrics Syst. Pharmacol, 2015. **4**: p. 189-192.
78. Chapman, K.L., et al., *Pharmaceutical toxicology: designing studies to reduce animal use, while maximizing human translation.* Regul Toxicol Pharmacol, 2013. **66**(1): p. 88-103.
79. Ajmera, I., et al., *The impact of mathematical modeling on the understanding of diabetes and related complications.* CPT Pharmacometrics Syst Pharmacol., 2013. **Jul 10**(2).
80. Langley, G.R., et al., *Towards a 21st-century roadmap for biomedical research and drug discovery: consensus report and recommendations.* Drug Discov Today, 2017. **22**(2): p. 327-339.