Emerging Technologies and Their Impact on Regulatory Science

Elke Anklam[m], Martin Iain Bahl[b], Robert Ball[a], Richard D. Beger[a], Jonathan Cohen[a], Suzanne Fitzpatrick[a], Philippe Girard[i], Blanka Halamoda-Kenzaoui[m], Denise Hinton[a], Akihiko Hirose[e], Arnd Hoeveler[m], Masamitsu Honma[e], Marta Hugas[c], Seichi Ishida[e], George E N Kass[c], Hajime Kojima[e], Ira Krefting[a], Serguei Liachenko[a], Yan Liu[g], Shane Masters[a], Uwe Marx[h], Timothy McCarthy[j], Tim Mercer[l], Anil Patri[a], Carmen Pelaez[Y], Munir Pirmohamed[aa], Stefan Platz[w], Alexandre J. S. Ribeiro[a], Joseph V Rodricks[o], Ivan Rusyn[p], Reza M Salek[r], Reinhilde Schoonjans[c], Primal Silva[k], Clive N. Svendsen[z], Susan Sumner[x], Kyung Sung[a], Danilo Tagle[v], Li Tong[d], Weida Tong[a], Janny van den Eijnden-van-Raaij[n], Neil Vary[k], Tao Wang[q], John Waterton[f], May Wang[d], Hairuo Wen[s], David Wishart[t], Yinyin Yuan[u], William Slikker[a].

a. U.S. Food and Drug Administration
b. National Food Institute, Technical University, Denmark.
c. European Food Safety Authority, EU
d. Universities of Georgia Tech and Emory, USA
e. National Institute of Health Sciences, Japan
f. University of Manchester and Bioxydyn Ltd, United Kingdom
g. Median Technologies, France
h. TissUse, Germany
i. Swissmedic, Switzerland
j. Pfizer, USA
k. Canadian Food Inspection Agency, Canada
l. Garvan Institute, Australia
m. Joint Research Centre, EU
n. Dutch Organ-on-Chip Consortium, The Netherlands
o. Ramboll, USA
p. Texas A&M University, USA
q. National Medical Products Administration, China r. International Agency for Research on Cancer, France

s. National Institutes for Food and Drug Control, China

t. University of Alberta, Canada

u. The Institute of Cancer Research, United Kingdom

v. National Institutes of Health, United States

w. AstraZeneca, United Kingdom

x. University of North Carolina, Chapel Hill, USA

y. Spanish Council for Scientific Research, Spain

z. Cedars-Sinai Medical Center, USA

aa. University of Liverpool, United Kingdom

Running Title: Emerging Technologies and Regulatory Science

Corresponding author: William Slikker, Jr., National Center for Toxicological Research/US-FDA, 3900 NCTR Rd., Jefferson, AR, 72079, william.slikker@fda.hhs.gov

Abstract:

There is an evolution and increasing need for the utilization of emerging cellular, molecular and *in silico* technologies and novel approaches for safety assessment of food, drugs, and personal care products. Convergence of these emerging technologies is also enabling rapid advances and approaches that may impact regulatory decisions and approvals. Although the development of emerging technologies may allow rapid advances in regulatory decision making, there is concern that these new technologies have not been thoroughly evaluated to determine if they are ready for regulatory application, singularly or in combinations. The magnitude of these combined technical advances may outpace the ability to assess fit for purpose and to allow routine application of these new methods for regulatory purposes. There is a need to develop strategies to evaluate the new technologies to determine which ones are ready for regulatory use. The opportunity to apply these potentially faster, more accurate, and cost-effective approaches remains an important goal to facilitate their incorporation into regulatory use. However, without a clear strategy to evaluate emerging technologies rapidly and appropriately, the value of these efforts may go unrecognized or may take longer. It is important for the regulatory science field to keep up with the research in these technically advanced areas and to understand the science behind these new approaches. The regulatory field must understand the critical quality attributes of these novel approaches and learn from each other's experience so that workforces can be

2

trained to prepare for emerging global regulatory challenges. Moreover, it is essential that the regulatory community must work with the technology developers to harness collective capabilities towards developing a strategy for evaluation of these new and novel assessment tools.

Keywords: Emerging Technologies, Biomarkers, Regulatory Science, Risk Assessment, Bioimaging, Bioinformatics

Impact Statement:

Emerging technologies will play a major role in regulatory science in the future. One could argue that there has been an evolution of use and incorporation of new approaches from the very beginning of the safety assessment process. As the pace of development of novel approaches escalates, it is evident that assessment of the readiness for these new approaches to be incorporated into the assessment process is necessary. By examining the arears of Artificial Intelligence and Machine Learning; Omics, Biomarkers, and Precision Medicine; Microphysiological Systems and Stem Cells; Bioimaging and the Microbiome, clear examples as to how to assess the reproducibility, reliability and robustness of these new technologies have been revealed. In a group movement, there is a call for product developers, regulators, and academic researchers to work together to develop strategies to verify the utility of these novel approaches to predict impact on human health.

Introduction

The organizing committee of the GSRS20 recruited a world-class set of authors carefully selected to address the theme of Emerging Technologies and Their Impact on Regulatory Science. Although the development of emerging technologies may allow rapid advances in regulatory decision making, there is concern that these new technologies have not been thoroughly evaluated to determine if they are ready for regulatory application. The number and diversity of these alternate approaches may make it challenging to determine whether the approach is ready for routine use in the regulatory environment. There is a need to develop strategies to evaluate the new technologies to determine their reliability, reproducibility and robustness as applied to regulatory decision making. The new approaches need to be evaluated to determine if these potentially faster, more accurate, and cost-effective approaches are ready for regulatory use. A clear strategy needs to be developed to evaluate emerging technologies rapidly and appropriately. In addition, regulatory scientists need to work with the novel approaches to understand the strengths and limitations of these new approaches. The regulatory field must set acceptable quality standards for these novel approaches and share this information with others working in the field of regulatory science. It is critical that the regulatory community work with the technology developers to harness the full value of the new technology and develop a strategy for evaluation of these novel assessment approaches.

The carefully selected authors at the Global Summit on Regulatory Science (GSRS20) provide a full understanding of the scope and status of global accomplishment in the application of emerging technologies to regulatory science. The GSRS20 organizing committee under the leadership of the Global Coalition for Regulatory Science Research (GCRSR) assembled an outstanding set of international science thought leaders to address the promise of emerging technologies and their application to regulatory science.

A series of introductory contributions from global research/regulatory leaders is provided to set the stage and introduce the needs for and contributions of emerging technologies and their impact on regulatory science. The first presentation is from the director of America's National Institutes of Health, Dr. Francis Collins, a proven world leader in the development of new technology and the importance of applying emerging technologies for improving public health. Dr. Collins' comments are followed by Dr. Bernhard Url, Executive Director of the European Food Safety Authority, European Union, a world leader in assessing safety. He emphasizes the three challenges: speed, complexity, preparedness in the safety assessments in the world's food supply. Next, Dr. Elke Anklam, principal advisor on life sciences at the Joint

4

Research Center of the European Commission provides comments about emerging technologies and their huge potential impact on the risk assessment of drugs and chemicals. From the National Institute of Health Sciences, Japan, Dr. Masamitsu Honma, deputy director general, addresses the regulatory sciences from his perspective in Asia. Following that, Dr. Margaret Hamburg, foreign secretary for America's National Academy of Medicine, provides some historical background concerning the regulatory sciences and the Global Summit on Regulatory Science.

Next to provide comment is RADM Denise Hinton, the FDA Chief Scientist. She describes FDA's regulatory science and innovation initiatives. Dr. Primal Silva, the Chief Science Operating Officer at the Canadian Food Inspection Agency, provides his perspective on technologies and regulatory science as does Dr. Anand Shah, the former Deputy Commissioner for Medical and Scientific Affairs at the U.S. FDA. Dr. George Kass, from the European Food Safety Authority, introduces the developments in regulatory science in food safety from a European Union perspective.

Following these opening addresses, there are several presentations within the topic of Track A, Artificial Intelligence for Drug Development and Research Assessment. This segment is introduced by Weida Tong from NCTR/FDA and Arnd Hoeveler from the Joint Research Centre of the European Commission (EC-JRC). Next, Track B, focused on OMICS, Biomarkers, and Precision Medicine, introduced by Neil Vary and Primal Silva from the Canadian Food Inspection Agency and Richard Beger NCTR/FDA and Susan Sumner UNC, Chapel Hill. Track C, focused on Microphysiological Systems and Stem Cells, is introduced by William Slikker, Jr., NCTR/FDA, and Elke Anklam from EC-JRC. Track D is focused on bio-imaging research in regulatory science and is described by Drs. Serguei Liachenko from the NCTR/FDA and John Waterton from the University of Manchester, U.K. Finally, the study of the microbiome in regulatory science is described by George Kass and Reinhilde Schoonjans, from the European Food Safety Authority.

Global Thought Leaders' Opening Contributions

Francis S. Collins, MD, PhD, Director, National Institutes of Health, United States

This year's summit focuses on emerging technologies and their application to regulatory science. I want to give special thanks to all the presenters, and to Dr. Slikker, Dr. Tagle, and all of the

others who have worked hard to put this together, even in the face of the global COVID-19 pandemic and the need of carrying this out in a virtual format.

Emerging technologies are playing critical roles in the development of new approaches to address the safety of foods, drugs, and healthcare products. And their importance this year—and our collective charge to ensure their responsible application in regulatory science—is certainly highlighted by what's happening with COVID-19. This pandemic has taken on a remarkable priority in all our lives and in our work and it is why we're having this global summit in a virtual workspace, instead of being all together in one place.

NIH is proud to have played a role in the development of all the emerging technologies that are being featured in this summit. And I think these technologies, which are finding their way into all kinds of applications, are some of the greatest hits that I have seen developed during my more than a quarter century at NIH. So, it is highly appropriate that this summit focuses on trying make sure we are doing everything possible to advance such technologies and to use regulatory science to implement them in ways that that are safe and effective for the public.

Those technologies that we're talking about include microphysiological systems, otherwise known as tissue chips, that are being used for preclinical safety and efficacy studies, as well as toxicology studies, NIH is a world leader in developing this technology and has worked closely with FDA along the way.

A second area is artificial intelligence and machine learning. Everybody is busy figuring out ways to apply these technologies. We in the life sciences are finding some remarkable moments of opportunity there, including in drug development.

Thirdly, we have -omics, all of those -omics: genomics, transcriptomics, metabolomics, and so on. These technologies have many valuable applications, including for the discovery of biomarkers for advancing precision medicine and for gaining insights into details of how cells do what they do, how disease happens, and what we can do about it.

We also need to acknowledge the role of the microbiome in health and disease. Increasingly, we are recognizing that we humans are not an organism, we're a superorganism. So, we must think, in many instances, about interactions between ourselves and the microbes that live on us and in us, which can both help us stay healthy or, if things get out of balance, can cause illness. Obviously, that has important implications for regulatory science as well.

All of these things are a good fit for what we are asking you all to address during this virtual summit. It does demand, if you're getting into questions about regulatory decision-making, rigorous assessments of what we know—and what we don't know—may take us a bit beyond the traditional academic lab. I'm glad that we're having this gathering that mixes various perspectives together in useful ways. It's not just about mixing representatives of particular scientific communities, it's about countries as well. Certainly, the effort to try to harmonize regulatory decisions needs to be global and to depend upon close interactions between research funding and regulatory agencies.

Thinking back about the efforts that we have conducted since I've been an NIH director, which is now a little over 11 years, I'm particularly pleased that we were able, early on, to set up an NIH-FDA Leadership Council. This has provided an opportunity for the senior leadership of NIH and FDA to identify areas in which we could work together particularly effectively. A past example of this is what's been done with tissue chips. A current example is the development of next generation sequencing tests, and, of course, artificial intelligence and machine learning are in that space as well. We are counting on that leadership council to continue to be a valuable forum for figuring out ways that we can work together. Frankly, I think this summit will provide some valuable ideas for our leadership council about new challenges that we ought to take on or about existing areas that we might steer in a somewhat different direction. We're counting on learning from this gathering.

Coming back to COVID-19, certainly all of the emerging technologies that you're going to be talking about have opportunities to provide hope in these pandemic times. By advancing the cause of science, including regulatory science, we have opportunities to find real solutions at a pace that needs to be as rapid as it possibly can. We have this wonderful relationship with the FDA, along with industry partners and other government agencies, that is called ACTIV, which stands for Accelerating COVID-19 Therapeutic Interventions and Vaccines. That effort, started in April 2020, has accelerated the pace of designing master protocols, identifying clinical trial capacity, focusing on vaccines and therapeutics, and focusing also on preclinical efforts in ways that have never been done before. In the past, you might have contemplated putting together such a public-private partnership in a two-year period. But ACTIV was put together in about two weeks and continues to be a remarkable contributor to the fact that we are as far along as we are now, in terms of testing out and prioritizing therapeutics, vaccines, and other approaches. Certainly, FDA has been a critical part of that whole enterprise.

In the diagnostic arena, we also have very important issues. We recognize that testing for COVID-19, while it's come a long way, would be more advantageous if we had even more capabilities for doing point-of-care testing. That's what the RADx program, standing for Rapid Acceleration of Diagnostics, aims to do by bringing new platforms forward, getting them validated, going to FDA, and seeking Emergency Use Authorization (EUA) approval of those so they can be deployed. Of course, we count on FDA to be rigorous in those analyses, so that we're sure that what we're offering the public happens quickly, but also in a way that can be trusted.

The need for the kind of conversation we're having at this GSRS20 summit can hardly be overemphasized. In the midst of everything else that's swirling around us, I really appreciate people taking the time to come and make presentations, and others to listen carefully and engage in discussion and to see if there are specific actions coming out of this that could be a useful contribution to the critical path forward.

That means celebrating accomplishments, but also looking closely and honestly at weaknesses to figure out ways that we could collectively address those. So, all of that's on the table. Again, I am sorry not to be able to look out over a sea of faces of people who are going to be engaged in this enterprise. I will have to satisfy myself by knowing you're there and being confident with the leadership that has put this agenda together. Some really good science will get presented and talked about, and some actions will be identified. I look forward to hearing what actions are moving forward from the leadership.

So, to all the presenters, organizers, and participants, I want to say thank you to everybody and may you have a wonderfully productive GSRS20.

Bernhard Url, PhD Executive Director, European Food Safety Authority (EFSA)

EFSA is tasked to provide scientific advice to the EU institutions and Member States about risks in the food and feed chain. EFSA deals with animal health, animal welfare, plant health, nutrition and maybe in the future with sustainability.

I envision common challenges (and opportunities) in EFSA's core business of finding, selecting, appraising, and integrating diverse streams of evidence to answer risk questions. I identify three challenges: speed, complexity, preparedness:

- **Speed:** the pace of innovation in the outside world is faster than innovation in our work procedures, which means there is a lag in method developments, up to date safety dossiers and a lack of speed in delivering scientific advice.
- **Complexity:** the increasing complexity of the food system, which relates to the physical material that is moved across the globe, and about the exponential growth of data and evidence. Society demands to be more protected while at the same time not trusting experts that much.
- **Preparedness:** how can science-based bodies master the paradox to have process stability to deliver efficiently *and* have enough fluidity to absorb outside complexities?

On all three challenges the work of the GCRSR is valuable as it aims to develop methods which use 21st century science for regulatory purposes. This touches data sciences, exposure sciences as well as new insights in toxicology and epidemiology.

I propose to further deepen collaboration between EFSA and the GCRSR towards a regulatory ecosystems approach. An ecosystem is an agile and co-evolving community of diverse actors who constantly sophisticate their collaboration, and sometimes also competition, to create more value. I conclude by stating that I am convinced that this deep collaboration will help organisations to stay relevant, now as well as in the future.

Elke Anklam, PhD, Principal Advisor at the Joint Research Centre, European Commission (JRC/EU)

The JRC is the European Commission's in-house scientific service and supports EU policymakers through providing multidisciplinary scientific and technical support.

One of the focus of the JRC is on the harmonization of analytical and toxicological methodologies that can be used, for example, in risk assessment of chemicals including nanomaterials. Therefore, the JRC is a proud member of the Global Coalition for Regulatory Science Research. Together with EFSA and the European Medical Agency (EMA), the JRC represents the European Union in this important global coalition.

The annual global summits organized by the Global Coalition for Regulatory Science and Research aim to provide a platform for regulators, policymakers, and scientists to exchange views on the innovative technologies, methods, and regulatory assessments.
The JRC had the honor of hosting the ninth annual global summit that took place in September

2019 in Stresa, Italy. The topic of GSRS19 was on the scientific and regulatory challenges related to nanotechnology and nano plastics [1].

Emerging technologies have amongst others a huge potential in the risk assessment of drugs and chemicals. It is important to understand amongst other, the interaction of living cells and their environment. I stress that there is an urgent need to develop and harmonize strategies and to be faster in the evaluation of new technologies and I conclude that is only possible when working together at a global level. Therefore, I look forward to the contributions of the renowned scientists of GSRS20.

Masamitsu Honma, PhD, Deputy Director General, National Institute of Health Sciences, Japan

I will describe how regulatory science influences the medical field in Japan. Regulatory science has contributed new developments in the prevention, diagnosis, and treatment of diseases and has established a system for facilitating the practical use of pharmaceuticals and medical devices as quickly as possible. This has promoted life innovation (i.e., the realization of a healthy and long-lived society through innovative medicines and medical devices originating in Japan). Four institutions promote regulatory science in Japan. The Ministry of Health, Labour and Welfare (MHLW) decides and organizes basic policy. The Pharmaceuticals and Medical Devices Agency (PMDA) is in charge of drug review and consultation. The National Institute of Health Sciences (NIHS) is responsible for developing official tests and guidelines as well as conducting basic research. Finally, the Japan Agency for Medical Research and Development (AMED) provides funding to facilitate research and development.

In 2018, the PMDA established a Regulatory Science Center to act as a command center; this center plays a critical central role in the incorporation of innovation in the regulatory system. This has led to the utilization of clinical trial data and electronic health records for advanced reviews and safety measures and has promoted innovative approaches to develop advancements in therapy and technologies. The Regulatory Science Center comprises of three offices: The Office of Medical Informatics and Epidemiology, Office of Research Promotion, and Office of Advanced Evaluation with Electronic Data. These closely work with other offices to review new drug reviews and safety measures such as advanced analysis of clinical trial data based on modeling and simulation, as well as pharmaco-epidemiological investigations using real-world data.

One innovative approach that the PMDA has been focused on is horizon scanning, which is aimed at detecting signs of emerging technology in very early stages. Previously, the PMDA did not assess emerging technology until after it was applied to product development. With the emergence of innovations, PMDA is now looking a step ahead and has prepared the means to assess new technologies properly.

I will introduce emerging technologies used by the NIHS to assess the quality, safety, and efficacy of pharmaceuticals, foods, and numerous chemicals. We mainly focused on four technologies:

- *In silico*/deep learning/artificial intelligence (AI)

- Omics; toxicogenomics technology

- Microphysiological system (MPS)/body-on-chip

- Desorption electrospray ionization–mass spectrometry (DESI-MS)

The NIHS has started to develop a large chemical safety database and AI platform to support efficient and reliable human safety assessment of pharmaceuticals, food, and household chemicals, where large-scale and reliable toxicity test data of the NIHS are integrated with expert knowledge beyond the regulatory framework. The aim is to support the introduction of a reliable management standard for pharmaceuticals, to prevent side effects, and to set safety evaluation criteria for food and household chemicals to avoid overdose and exposure. A success of this project has been the development of a chemical mutagenicity prediction model using quantitative structure-activity relationship (QSAR) and AI [2].

Research on toxicogenomics has been promoted since 2000. For example, the Percellome project was initiated to develop a comprehensive gene network for mechanism-based predictive toxicology. This project focuses on building and maintaining the Percellome database for single-dose toxicity; this is one of the largest transcriptome databases accessible *via* the Internet.

An MPS is a chip that introduces cells onto a microfluidic device with the aim of reproducing tissues and physiological functions that are difficult to obtain with conventional cell cultures and experimental methods. An MPS enables a more accurate evaluation of the effects and toxicity of drugs on humans.

DESI-MS is an ambient ionization technique that can be coupled with mass spectrometry for chemical analysis of samples under atmospheric conditions. The charged droplets generated by the electrospray impact the surface, where the analyte is dissolved into the electrically charged droplets. The secondary droplets ejected from the surface are subsequently collected into an ion-transfer tube and are analyzed by mass spectrometry. This technique allows the distribution of the administered drugs and metabolites in tissue sections to be visualized through the detection of individual mass images [3].

Margaret Hamburg, MD, Foreign Secretary, National Academy of Medicine

It's a great pleasure to help launch the 10th Global Summit on Regulatory Science. This international conference represents a very special event bringing together scientists from government, industry, and academic research from around the world to explore critical advances in science, technology, and innovation and to strengthen and extend partnerships needed to enhance translation of these scientific advances into regulatory applications; applications that result in much-needed products, and applications that can be applied within a global context. This year's theme, "Emerging Technologies," and their application to regulatory science is exciting and exceedingly timely.

Ten years ago, when this global summit first began, I was serving as Commissioner of the U.S. Food and Drug Administration. There I had to good fortune to work closely with Bill Slikker who's been the vision and energy behind this summit from the very beginning. I was very excited by the possibility of what a summit such as this could offer, but whether the summit idea would last one year or a decade or more was uncertain. We suspected that such a convening could have real and enduring value and I think there's no question about that now. So, I certainly want to thank Bill and the team that has put together this meeting, as well as all of those who have helped to develop and to host these meetings over the past 10 years. I also want to welcome and thank all of you for participating in this global summit. Your work matters.

At FDA I came to understand how truly vital regulatory science is to the overall scientific enterprise. It is what enables us to harness the power of science and technology in the service of people and of progress. It speeds innovation, streamlines research and development, improves regulatory decision-making, and strengthens our ability to get the safe and effective products necessary to address unmet public health needs and medical care needs.

To be effective, regulatory science must ground itself on several things including strong science, partnerships, and the ability to work across disciplines, sectors, and borders. This regulatory summit has from the beginning been committed to doing just those things. It has offered a really valuable and somewhat unique platform where regulators, policymakers, and bench scientists from a wide range of countries can come together to discuss how to develop, how to validate, and apply innovative methodologies and approaches for product development and regulatory assessments within their own countries, as well as harmonizing and aligning strategies to a shared collaborative global engagement.

Now in the midst of a global pandemic, I think we all appreciate the value of this more profoundly than ever. In fact, I think the whole world is looking to regulatory science to facilitate the swift but scientifically robust development, review, approval, and availability of essential medical products from personal protective gear to diagnostics, drugs, and vaccines. Too often regulation had been viewed as a barrier to progress but in fact it represents a crucial partner in achieving the shared goals of developing innovative, safe, and effective products as effectively and efficiently as possible. This is what all stakeholders want and expect. As desperate as we are for drugs, vaccines, and new improved diagnostics for COVID-19, the public and patients along with the medical community, policy makers, and government leaders all must have trust and confidence in the process and the products.

In many ways regulatory science represents the bridge to both translate the science into the products we so desperately need and to ensure a robust process based on science. Rigorous science does not have to mean rigid, inflexible, and slow. And speed does not have to mean cutting corners. In fact, applying innovative regulatory science enables us to apply more modern and adaptable tools to the R&D and regulatory review, approval, and oversight process including such things as biomarkers, predicted toxicology, microphysiological systems for drug development, nanotechnology-driven applications, innovative clinical trial designs, and artificial intelligence, and IT-driven strategies for product development, risk assessment, and regulatory monitoring.

Not surprisingly, I see that you'll be addressing many of these areas of science in this meeting as you think about emerging technologies and their applications to regulatory science. We are living in a time when the advances in science and technology are unfolding with unprecedented speed. Yet the challenges we need to address, with the best possible science, are also unprecedented and profound. That is why the work each of you does is so important and why coming together now to deepen understanding, gain new insights, and strengthen collaboration

13

will enable us to better define and implement the best science-driven regulatory practices that can support the public health imperatives of a global pandemic such as COVID-19, but can also assure systems in which the best discoveries in science can be translated as quickly and appropriately as possible to deliver the kinds of the innovative, safe, and effective products that patients and consumers expect and deserve each and every day.

Denise M. Hinton, RADM, Chief Scientist, FDA, USA

The 2020 Global Summit on Regulatory Science highlights Emerging Technologies for Regulatory Application, a Global Perspective. The topic is certainly appropriate, given the current state of affairs - an unprecedented pandemic. And yet, never has there been a greater need for regulatory science to embrace and address the interconnection among people, animals, plants, and their shared environment - a timeless concept that we call One Health.

Globalization, new science, emerging technologies have been the accelerators that are bringing us together, literally and now virtually, to address both their inherent opportunities and their challenges. FDA has long valued collaboration in leveraging broad expertise to tackle a problem. It is even clearer what the confluence of these efforts has demanded we do if we want to successfully protect, promote, and advance the public health.

We must continue to find new ways of collaborating with all our stakeholders in the regulatory science enterprise. The current and potential regulatory science applications of exciting technologies highlighted in this global summit include, but are not limited to artificial intelligence, microphysiological systems, the microbiome, biomarkers, and precision medicine, all of which are evidence of this essential collaborative approach and action.

Just a decade ago, the Office of the Chief Scientist was established at FDA to forge many of the intramural and extramural[1,2] *(https://www.fda.gov/science-research/advancing-regulatory-science/centers-excellence-regulatory-science-and-innovation-cersis; https://www.fda.gov/science-research/advancing-regulatory-science/regulatory-science-extramural-research-and-development-projects)*collaborations that are now helping to advance these technologies, and to support a global network of partners who can help ensure the safety of our medical products and our food supply. For instance, FDA's cross-agency scientific work is focused on developing and fostering opportunities for promising innovative technologies with the goal of advancing new tools as well as new areas of science.

For example, FDA's Alternative Methods Working Group[3] is focused on supporting methods that are alternatives to traditional toxicity and efficacy testing that will be used across FDA's various product areas. This working group is serving as a catalyst to the development and potential application of alternative systems in vitro, in vivo, and in silico, and using systems toxicology and modeling to inform FDA decision-making and regulatory toxicology. To that end, the Alternative Methods Working Group has launched a webinar series[4] (*https://www.fda.gov/science-research/about-science-research-fda/fda-webinar-series-alternative-methods-showcasing-cutting-edge-technologies-disease-modeling*) for developers to share their cutting-edge technologies with FDA scientists.

Well-established collaborations with our stakeholders at home and abroad are enabling FDA to work expeditiously in support of the application of these new technologies to the development of therapeutics and vaccines for COVID-19. By leveraging nimble funding mechanisms like the OCS-led advancing regulatory science broad agency announcement (BAA), FDA has been able to solicit innovative ideas and approaches to evaluating FDA regulated products and build on long standing relationships with our collaborators.

At no time have the benefits of these existing partnerships with world class institutions been more caring than during this pandemic. For example, the Office of Counterterrorism and Emerging Threats, OCET in OCS, is using BAA-funded contracts to support FDA medical countermeasure research on Ebola and Zika.

One study, with Public Health England,[5] *(https://www.fda.gov/emergency-preparedness-and-response/mcm-regulatory-science/developing-toolkit-assess-efficacy-ebola-vaccines-and-therapeutics)* has been expanded to leverage technology used in developing a toolkit to assess efficacy of Ebola vaccines and therapeutics to gather vital information about COVID-19 infections. This project is generating reagents that are being shared with FDA scientists and external partners to support COVID-19 research. BAA project with Stanford University[6] *(https://www.fda.gov/emergency-preparedness-and-response/mcm-regulatory-science/survivor-studies-better-understanding-ebolas-after-effects-help-find-new-treatments)* that is evaluating the after-effects of Ebola in survivors, and how to more effectively treat these patient's chronic health problems has also been extended to aid the development of rapid diagnostics, therapeutics, and vaccines for COVID-19 and inform FDA evaluation of these products.

And, most recently, in partnership with National Institute of Allergy and Infectious Diseases, OCET has awarded a new BAA contract to the University of Liverpool[7]

(https://www.fda.gov/emergency-preparedness-and-response/mcm-regulatory-science/fda-and-global-partners-analyze-coronavirus-samples) who, together with a global consortium, will sequence and analyze samples from humans and animals to create profiles of various coronaviruses, including SARS CoV-2. The investigation will also use in vitro coronavirus models like organs-on-a-chip to help inform development and evaluation of medical countermeasures for COVID-19.example where we have built on existing collaborative research, FDA and its partners at NIH have recently modified a crowdsourcing application co-developed by FDA and the Johns Hopkins CERSI to enable the clinical community to share their COVID-19 treatment experiences. Our UCSF/Stanford CERSI has begun research on a rapid query model that is enabling us to examine COVID-19 questions using electronic health record data.

The "all-hands-on-deck" response to the COVID-19 pandemic has brought home, not only the criticality of worldwide collaboration and cooperation in combating this global scourge, but the recognition that the health of people, animals, and the ecosystem are interdependent. In 2019, FDA launched its One Health Platform,[8] *(https://www.fda.gov/animal-veterinary/animal-health-literacy/one-health-its-all-us)* to further public and global health and breakdown silos.

Working at the nexus of the three one health domains of human, animal, and environmental health, FDA has been encouraging the expansion of multi-disciplinary and inter sectorial networks and partnerships to improve our scientific knowledge with information-sharing, pooling of skills and resources, and monitoring and surveillance with the goal of generating more beneficial health outcomes.

NCTR, Jefferson Labs, is surveying SARS CoV-2 in wastewater, which is a research project that exemplifies the One Health approach as a complimentary tool for estimating the viral speed of COVID-19 in the central Arkansas area. The goals are to establish a method to extract and quantify SARS CoV-2 in wastewater samples; monitor the temporal dynamics of titer of SARS CoV-2 in the wastewater as a proxy for the presence and prevalence of COVID-19 in Arkansas communities; and estimate the number of clinical cases, based on the titers of SARS CoV-2 in the wastewater sample.

This kind of early detection and continuous monitoring of COVID in the community may help federal and local agencies respond more quickly to halt the spread of COVID-19 and decrease the burden of patient admissions in healthcare facilities in the event of future pandemic waves—a direct public health impact.

Let this pandemic, with its clear evidence of the interconnection of plants, animal, and human life serves as a plan of action for future public health emergencies. FDA's 2021 Focus Areas of Regulatory Science[9] *(2021 Advancing Regulatory Science at FDA: Focus Areas of Regulatory Science (FARS). 2020. https://www.fda.gov/media/145001/download)* reflects the need for a nimbler approach that will enable us to respond swiftly to the increasing rapid evolutions in science and technology.

In building a robust scientific infrastructure and training programs for scientists, FDA is laying the foundational elements to tackle innovations in our regulatory portfolio and public health preparedness and response. What is needed is a preparedness mindset and the global drive to develop and implement the interventions that will make us resilient in the face of challenges to come in our increasingly interconnected world.

Anand Shah, MD, former Deputy Commissioner of Medical and Scientific Affairs, FDA, USA

This year's global summit with a focus on emerging technologies and their impact on regulatory science emphasizes an issue that has always been important given how the relentless nature of scientific progress requires regulators to continually improve and innovate their regulatory processes to ensure the safety of consumer products.

However, this theme is particularly salient this year in 2020 given the ongoing COVID-19 pandemic with the virus infecting more than 23 million people and claiming the lives of over 800,000 individuals worldwide. Given the SARS-CoV-2 is a novel pathogen, we didn't have any medical products designed to diagnose or treat this specific virus when the outbreak first broke out.

Over the past eight months, the innovation from the scientific and medical communities has been simply tremendous. We've seen an array of timely and innovative tests and some highly accurate PCR-based diagnostics to test that use saliva to increase speed and convenience. In terms of therapeutics, we've seen an incredibly diverse pipeline of potential treatments emerge with hundreds of clinical trials initiated worldwide for antiviral drugs, immunomodulators, neutralizing antibodies, and more.

And of course, as the world works to chart a roadmap to recovery, there's been an increasing focus on COVID-19 vaccines with tens of thousands of individuals volunteering across the world

to participate in clinical trials over the past few months. These emerging technologies are exciting, highlight the ingenuity of the research community, and offer hope to patients and consumers that there's light at the end of the tunnel of this pandemic.

However, as regulators, the crisis and associated emerging medical technologies have created a unique series of challenges. During emergency situations, it's important for us to be responsive to the evolving public health need and be adaptive so that our regulatory decisions are made on the latest and highest quality evidence.

The public and medical providers have tremendous faith in the decisions of regulators. And it's important that we, in turn, demonstrate our commitment to safety and scientific integrity. Consider the case of COVID-19 vaccines. While we celebrate the unprecedented collaboration and scientific focus on developing these new products as public health officials, we're concerned by the reports of growing vaccine hesitancy in the population. Vaccines are foundational to public health and have played a critical role in reducing morbidity and mortality from infectious diseases for over 200 years.

For COVID-19, vaccines that are shown to be safe and effective through rigorous clinical testing can help us safely achieve herd immunity and return to normal life. It's important for the public to know that regulators are committed to following unwavering regulatory safeguards for COVID-19 vaccines.

Commissioner Hahn, Dr. Peter Marks, and I recently reaffirmed this commitment in an article in JAMA where we outline some of the key steps the agency has taken such as providing guidance to vaccine developers in committing to seek input from the FDA vaccine advisory committee to help ensure that the public is clear about FDA's expectations for data to support safety, effectiveness and that the required regulatory standards will be met, so they can be comfortable receiving any vaccine to prevent COVID-19.

While COVID-19 is our number one priority right now, it's important to recognize that the challenges and lessons for regulatory science during this pandemic are applicable to a whole host of other ongoing and forthcoming public health issues. As all attendees of this summit are aware, there's a continuing evolution of the utilization of emerging technologies and novel approaches for the safety assessment of food, drugs, and personal care products. Prominent examples include the application of genomics, proteomics, and metabolomic technologies. These technologies have served as the foundation for precision medicine and improve food

safety and traceback procedures.

For example, genomic analysis of pathogens is becoming increasingly common. This is a welcomed innovation as genomic analysis can help us improve surveillance for the public health threat of antimicrobial resistance, and support consumers by advancing the movement towards personalized medicine. However, as these technologies advance, regulators need to be informed and aware of the best practices for verifying the evidence and validating the performance of these products.

Another example is proteomics which has significant potential for improving food safety by enabling screenings for foodborne pathogens or allergens with high sensitivity and specificity. These advances would help improve the traceability and quality of the food system – a commitment the FDA recently affirmed in its blueprint for the new era of smarter food safety.

Other emerging technologies which will be featured at this year's conference include the use of artificial intelligence and bioinformatics in predictive systems and the use of MPS such as "organ-on-a-chip" technologies which can help bridge the gap between *in vitro* and *in vivo*. While each of these technologies offers tremendous potential, it's important that regulators possess the tools and capacity needed to systematically examine the strengths and weaknesses of each approach.

Furthermore, given the increasingly interconnected nature of today's world, it's also important for regulators to share experiences and best practices with one another so that we can harmonize our standards where possible and prepare our workforce to meet the challenges of emerging technologies. That's why forums like today's global summit on regulatory science are so important. By convening the world's leading researchers, regulators, and experts on these technologies, we can help activate our collective capabilities and support the development of a strategy for evaluating these novel assessment tools.

Primal Silva, PhD., Chief Science Operating Officer, Canadian Food Inspection Agency, Canada

I will focus on how the Canadian Food Inspection Agency (CFIA) is using emerging technologies to expand and enhance its capabilities in carrying out its regulatory function.

The CFIA is an organization in Canada that is tasked with doing many different functions. It has a mandate in terms of prevention and managing food safety risks in Canada. It also is tasked with the plant health and animal health responsibilities in the country, as well as contributing to the consumer protection and facilitating market access for Canada's food, animals, and plants. CFIA is the largest science-based regulatory agency in Canada and uses technology and well-trained scientists and expertise to carry out its function. It is supported by 13 laboratories spread across the country that use cutting-edge technologies to deliver on its broad mandate.

In terms of some of the technologies, I want to describe a new program in Canada, Innovative Solutions Canada, where government departments engage with small and medium-sized enterprises with innovative new ideas to see how they can help deliver government services. We have partnered with several companies using this program to bring about the best available science and technologies to help us carry out regulatory functions. For example, with the current COVID-19 pandemic, we are validating a ONETest CoronaVirus test that can test for the presence of the virus, not only in human samples, but across some of the animal species given the broad host range and the zoonotic nature of this virus.

- The other area where the CFIA has done a lot of work is in applying genomics for helping with regulating food, in terms of the whole genome sequencing that is now fully integrated into the testing for foodborne pathogens in the Agency as well as for antimicrobial resistance testing. Recently in Canada, we switched over from pulsed-field gel electrophoresis to whole genome sequencing technology to make the link between human foodborne pathogen samples and food isolates [4, 5]

- Similar advances have been made in the plant health area [6]. For example, we are using whole genomic sequencing to detect plant viruses of quarantine significance as way to reduce the time and improve precision in permitting release of new plant varieties in Canada. Likewise, new genomic technologies play a very big part in the animal health diagnostics e.g., chronic-wasting disease, rabies, avian influenza, African swine fever etc. These new technologies are often used in conjunction with the classical tests that are already established as the gold standard tests for diagnostics [7].

In terms of advancing regulatory application of genomics, it has been crucial that we work with international partners for the validation of these technologies. Addressing key questions such as - how do you standardize genomics-based results, or how do you apply genomics in the regulatory context for decision making? – are all important in this context.

I want to acknowledge the important role the Global Coalition for Regulatory Science Research has played in developing international consensus on scientific matters. As global regulators, we have a very important role in promoting science and evidence-based international standards and action.

George Kass, Ph.D., Lead Expert at the European Food Safety Authority, EU

I will address emerging technologies for regulatory applications with a focus on the EU food safety perspective. EFSA is the European Union (EU) agency that is responsible for providing independent scientific advice and for communicating on existing and emerging risks associated with the food chain in the EU[8]. Its advice forms the basis for European policies and legislation and its remit covers food and feed safety, nutrition, animal health and welfare, plant protection, and plant health. Furthermore, EFSA also considers, through environmental risk assessments, the possible impact of the food chain on the biodiversity of plant and animal habitats.

Many food and feed related products require a scientific assessment to evaluate their safety before they are authorized on the EU market. These so-called regulated products include substances used in food and feed (such as additives, enzymes, flavorings, and nutrient sources), novel foods, food contact materials and pesticides, genetically modified organisms, food-related processes, and processing aids, and these require a specific risk assessment to be performed; These products may be new or already on the market. In addition, EFSA is dealing with other types of substances that may be found unintentionally in the food chain, such as contaminants.

For regulated products, the type and amount of data provided to EFSA depends very much on the data requirements, the latter being specified either in the prevailing EU legislations or in the appropriate guidance documents developed by EFSA. Often the amount and type of information requested depends on the application domain and use levels but also follows EU legislation when it comes to the minimization and optimization of the use of animal data, the so-called 3R principles, in accordance with Directive 2010/63/EU. The type of data typically received by EFSA consists of *in vivo* and *in vitro* studies but also includes *in silico* data. The *in vivo* studies, which predominate, include studies on absorption, distribution, metabolism, and excretion (ADME), aiming at information on the behavior of a substance in a complex biological system, as well as toxicological studies to identify and characterize the potential toxicity of this substance. *In vitro* studies are also evaluated by EFSA, but these are mainly genotoxicity studies.

The development of new technologies, methodologies and tools for chemical risk assessment presents for an organization like EFSA both challenges and opportunities. The types of new methodologies that EFSA is facing are often referred to as New Approach Methodologies (NAMs) (European Chemicals Agency, 2016; *https://echa.europa.eu/documents/10162/22816069/scientific*). The overarching goal of NAMs is to enable the replacement of animal testing through a combination of predictive in silico models and in vitro assays. On the in-silico side, these can be tools to predict the physico-chemical properties of chemicals or the behavior of chemicals in biological systems. The latter include (Q)SAR and read-across tools as well as PBTK models to predict the kinetics of a chemical, across species and across biological systems (e.g., QIVIVE models to link external and internal exposure). Attempts to develop more sophisticated and elaborated systems such as such as virtual organ models and virtual organisms are also currently under development. Better performing *in vitro* tools are being developed at an equally rapid pace to overcome the limitations of 2D culture systems based on transformed cell lines. For instance, stem cell differentiation protocols are becoming mature to produce human cell models that not only are more biologically relevant than experimental animal-derived cells but also are amenable to greater complexity using in vivo-like architectures such as organoids, microphysiological systems on chip to mimic multi-organ systems or even the whole-body architecture.

An impact from many of these new methodologies is the quantity of data generated, the so-called big data, which can come from high throughput screening of multiple chemicals or from high content screening on individual substances. Also, whole genome sequencing produces a considerable amount of data as also other OMICs approaches do. With this challenge comes the need for new tools and skills to handle big data, not only for their collection, their curation, their storage, but, also, for their analysis. While progress is being made in these areas, it is encouraging to note that rapid developments are being seen at the level of artificial intelligence and machine learning at helping in the prediction of toxicity and in the risk assessment.

When assessing the regulatory landscape for food safety, it becomes evident that the great majority of biological and toxicological data requested by legislation involve *in vivo* studies that are expected to be conducted using the appropriate test guidelines of the Organization for Economic Co-operation and Development (OECD) and following the OECD principles of Good Laboratory Practice (GLP). In the case where EFSA develops its own sector-specific guidance documents, these equally describe the types of data and studies that should be performed to support an appropriate risk assessment. It is important to stress that the same requirements for following OECD test guidelines and the principles of GLP apply. However, where possible, new methodologies are referred to and incorporated, but under the condition that such methodologies

are sufficiently robust, reproducible, and accepted for regulatory decision making. EFSA's policy to update its guidance documents on a regular basis also provides an opportunity for new methodologies to be incorporated as they become accepted in the regulatory arena.

NAMs carry several advantages, including the use of human relevant models. They can produce mechanistic data, too, that informs the risk assessor on adverse outcome pathways and modes of action. Furthermore, the information generated can be much more human disease-relevant than from animal studies and can be produced at a pace and cost that support a more rapid risk assessment. Yet, we also see hurdles in their use in the regulatory context. The new methodologies need to be standardized, be reproducible and need to be validated in order to gain the confidence of the risk assessor. Moreover, they also need to be accepted by the wider community. For these reasons, EFSA is very much engaged with organizations that deal with the issues of promoting, developing and validating NAMs. We work very closely with the European Centre for the Validation of Alternative Methods (ECVAM) at the JRC but also with the European Chemicals Agency (ECHA). On the EU research and innovation front for regulatory sciences, it is also important to highlight the development of the new Horizon Europe framework program that will be launched in 2021 and that includes the co-creation of the Partnership for the Assessment of Risks from Chemicals (PARC) *(https://ec.europa.eu/info/horizon-europe/european-partnerships-horizon-europe_en)*. Together with still ongoing Horizon 2020 programs such as EUToxRisk *(https://www.eu-toxrisk.eu/)*, such initiatives are expected to help promote the uptake and use of NAMs in regulatory sciences. In addition to its engagement with such initiatives, EFSA also has its own vehicle of promoting and funding research into areas directly pertinent to new methodologies, such as its newly created Science Studies and Project Identification & Development Office (SPIDO) under which, among others, NAM case studies have been launched. At an international level beyond the EU, EFSA collaborates with the World Health Organization, with OECD, with the Global Coalition for Regulatory Science Research, and other food safety agencies outside the EU. EFSA is also actively involved in programs such as the international government-to-government initiative, Accelerating the Pace of Chemical Risk Assessment (APCRA), and has established its own platforms such as the International Liaison Group for Methods on Risk Assessment of Chemicals in Food (ILMERAC) to interact with various organizations in its ambition to contribute to the harmonization of chemical risk assessment to support food and feed safety.

Christopher Austin, PhD, former Director of the National Center for Advancing Translational Sciences (NCATS) at the National Institutes of Health, USA

Of the 27 institutes of the NIH, NCATS is the newest of them, and our job is closely related to the purpose of this meeting focused on emerging technologies. That is, to develop new ways, new methods, new technologies, new paradigms to make the process of developing and deploying interventions that improve human health better, faster, and we hope, cheaper. So, we work on the problem of how therapeutics and diagnostics and medical procedures are developed and improve that efficiency, and regulatory science is big part of that.

So, how did we define translational science? Like any other science, it is a field of investigation that seeks to understand general principles – in this case the underlying principles of the translational process – with the ultimate goal of prediction. I think about regulatory science as largely a subset of translational science.

The problems we work on are the major rate-limiting steps to translational efficiency. A lot of you will find these familiar, and ones that you run into every day, and bedevil every therapeutic area. And you'll notice, if you look down this list, a very large number of them are regulatory issues. Not all, of course, because many interventions are not subject to regulatory approval, behavior interventions as an example.

So, the first thing I want to tell you about, which I hope you find useful in your own work, is something that we did with a group at the National Academy of Medicine a couple of years ago, to address a problem that is familiar to all of you: that many stakeholders, if not most, are unaware of the complexity of the process by which interventions that improve human health are developed. We felt that there was a need for an accurate portrayal of what is involved, not only as an educational tool, but also for use by translational and regulatory scientists to determine where they should apply their efforts, based on knowledge of the most problematic steps in terms of failure, time, or cost. We published process maps for small molecules and biologics in a couple papers a couple years ago that you might want to look at.

The group took as its starting place the most commonly used portrayal of therapeutic development, the linear chevron diagram. It is unfortunately a terribly inaccurate and misleading portrayal since it makes the process appear straightforward – drug development in 6 easy steps! So, we broke out each of the stages of the chevron into a "neighborhood" of dozens of interconnected and recursive steps. Since the publications of the map in static form in 2018, we have converted it into an interactive electronic form that you can find on the NCATS website (https://ncats.nih.gov/translation/maps), with the ability to zoom in and out and see where the

problem areas are and go to resources which might help you through them.

NCATS's longest running regulatory science initiative is the Toxicology in the 21st Century (Tox21) program, a collaboration among NCATS, FDA, EPA, and the National Toxicology Program which is developing innovative test systems, data, and algorithms that better predict human toxicity. Current goals of the program include implementing technologies that overcome some of the traditional limitations of *in vitro* testing systems, such as metabolic capability, and testing systems that didn't exist when we started Tox21, such as Induced Pluripotent Stem Cells (iPS) cells, 3D cellular organoids, and tissue chips, utilizing transcriptional readouts in addition to biochemical or imaging ones, and working via the adverse outcome pathway paradigm that I think you're all familiar with.

In addition to generating unprecedented amounts of in vitro screening and follow-up mechanistic and organ-based data, we are curating legacy data from the NTP and EPA, integrating it with Tox21 data, and making all data and analyses publicly available. We take great care in validating assay performance and compound identity and purity and are beginning to utilize these data in physiologically based pharmacokinetic (PBPK) modeling.

Complementing Tox21's systems biology approach are efforts to develop more predictive translational models and assays. At NCATS we utilized and support development of the full range of human cell-based assay platforms. We're all familiar with the traditional 2D cell line culture systems that many of us have spent our careers doing, but in the last ten years, it has become possible to test the activities of compounds in human primary cells, iPSC-derived cells, multicellular spheroids, multi-cell type organoids, 3D printed tissues, or microphysiological systems, otherwise known as MPS, tissue chips, or organs-on-a-chip. Each system is useful and complements the others, given the inverse relationship of throughput and physiological complexity. At one end of the spectrum, cell line based small volume multi-well (384- or 1536-well plate) HTS assays have low physiological complexity but can test up to hundreds of thousands of compounds a day. At the other extreme, tissue chips can be very physiologically complex can test only a few compounds a day.

Microphysiological systems (MPS) merit particular mention since they are the newest platform for compound testing so may be less familiar. This field has developed over just the last ten years, through the leadership of NCATS and a number of other institutes at NIH, DARPA, and FDA. MPS are multi-human cell type bioreactors that mimic the structure and function of human tissues. MPS have been developed for over a dozen individual human organs as well as

linked arrays of up to 10 different human organs. MPS are being used to study normal and disease physiology as well as to test for effects of xenobiotics on human tissues, currently as a complement to, but we hope eventually in partial replacement of, animal testing.

Population clinical data from humans is also a rich source of regulatory science-relevant insights, as FDA's Sentinel Initiative and many epidemiological studies have shown. However, there has to date not been an appropriately scaled, nationwide, publicly accessible resource of EHR data that is broadly representative of U.S. population demographics. NCATS and its Clinical and Translational Science Award (CTSA) grantees began creating the foundation of such a resource in 2018, and when the urgent need for such data became evident in the early days of the COVID-19 pandemic, the platform was rapidly instantiated to gather, harmonize, and make available in a secure privacy-protected federal enclave longitudinal EHR data on millions of patients with COVID-19 symptoms and diagnoses.

The National COVID Cohort Collaborative (N3C) currently has information on hundreds of thousands of patients from academic health centers across the country and will have approximately six million COVID-19 patients' records by early 2021. The purpose is to transform the clinical information that's in those electronic health records into a format that can be queried by individual investigators and have artificial intelligence or machine learning programs applied to them in ways that are impossible with current federated methods. And the kinds of things that one would be able to ask here are there -- what are the risk factors that indicate a better or worse prognosis? What are all the treatments that have been used for these patients that might make patients better or worse? The N3C (https://covid.cd2h.org/) will launch shortly and you can apply for access if you're a qualified investigator, with projects approved though a data access committee. Though the platform is focused on COVID-19 for now, it is completely generic in methodology, so once COVID-19 is over we hope to expand the platform's purview to other diseases and regulatory science applications.

Lastly, I want to tell you about CURE-ID (https://cure.ncats.io/), a different clinical data resource that I hope many of you will use and contribute to. Two years ago, scientists at NCATS began working with colleagues at the FDA to develop a mobile application that would allow health workers, particularly in the more remote parts of the country or the world, to share their clinical experiences using approved drugs to treat diseases other than the ones for which they're regulatorily indicated on the label. The project was initially focused on the neglected tropical diseases of the developing world. CURE-ID was launched in December 2019, and of course

later that month the first cases of COVID-19 were reported. A COVID-19 module was immediately added to CURE-ID, and rapidly became an invaluable resource for clinician experience sharing, particularly in the difficult early days of the pandemic. The platform currently covers over 100 infectious diseases, and we anticipate that it will be increasingly useful for both clinical and regulatory science applications.

Track A: Artificial Intelligence (AI) and Machine Learning (ML) for Regulatory Science Research, Risk Assessment and Public Health (Weida Tong, PhD, National Center for Toxicological Research/FDA, USA and Arnd Hoeveler, PhD, Joint Research Centre, European Commission, EU

Artificial Intelligence (AI) has impacted a board range of scientific disciplines and played an increasing role in regulatory science, risk assessment and public health. AI is a broad concept of training machines to think and behave like humans. It consists of a wide range of statistical and machinal leaning approaches to learn from the existing data/information to predict future outcomes. Some new AI methodologies such as deep learning has advanced algorithms and accuracy to extract complex patterns from new data streams and formats (e.g., image and graphic data) which has been increasingly used in the regulatory context. Data connectivity, computational resource and new/advanced algorithms fuel the rise of AI, which give new insight to underlying mechanisms and biology, thus significant for regulatory application. Although the concept of AI was introduced back in the 1950s and has been actively pursued in the research community, its critical role in regulatory application concerning food and drug safety, has yet to be been realized. Understanding that AI offers both opportunities and challenges to global regulatory agencies with questions such as (1) how to assess and evaluate AI-based products and (2) how to develop and implement AI-based application to improve the agencies' functions, GSRS20 hosted a special track to discuss the basic concept and methodologies of AI applied in regulatory science, risk assessment and public health with real-world examples. Particularly, the current thinking and on-going efforts with various agencies in applying AI in these fields were discussed with a specific focus on drug and food safety, clinical application (e.g., prognosis and diagnosis), precision medicine, biomarker development, natural language processing (NLP) for regulatory documents and regulatory science application.

One of the important questions raised by many presenters is about where the highest potential for the application of AI in regulatory agency is. It was well acknowledged that AI based NLP could

play a major role in extracting useful information from both the regulatory documents and literature data to facilitate regulatory science research and improve the product review process. It is a fact that one of the most important AI advancements in the recent years is in the area of NLP. The early effort in this area treated a document as a "bag-of-words", where the context of the document is pronounced based on the frequency of words. Recognizing that the context of a document is better represented by the co-occurrence of multiple words, the Probabilistic Graphic Models were introduced and have gained attraction in the community to understand "topics" of a document. However, more recently, the advanced AI in NLP offers an unprecedented opportunity to analyze the entire sentence, beyond words, by using language models to perform a broad range of NLP tasks such as Questions/Answers, Sentimental Analysis, Information Retrieval, Text Classification, etc. Dr. Henry Kautz from National Science Foundation presented examples for analysis of social media data for foodborne illness surveillance and illness prediction. The role of NLP was also discussed for drug target safety assessment by Dr. Stefan Platz from AstraZeneca as one of five pillars of AI for safety assessment – that is right target - target safety assessment, (2) right molecule – lead optimization chemical toxicity, (3) right safety – toxicological study, (4) right tissue – digital pathology, and (5) right patient – translational safety.

The regulatory agencies thorough out the world have routinely generated a variety of documents during regulated product review and decision-making, which has led to a large inventory of review documentation with a broad array of information. For example, Guidance documents, drug labeling, and FDA Adverse Event Reporting System (FAERS) are a part of the regulatory documents made available by U.S. Food and Drug Administration submissions (FDA). Meanwhile, the regulatory process also requires curating literature data for science-based decision-making. Unfortunately, most of these tasks are still a manually conducted process, which is time-consuming and labor-intensive, thus resulting in longer review times. Dr. Robert Ball from U.S. Food and Drug Administration (FDA) emphasized AI-based NLP as an important piece of modernization for the FDA Sentinel system. Both Dr. Philippe Girard from Swissmedic of Switzerland and Dr. Blanka Halamoda Kenzoui from EC-JRC provided specific examples on improving regulatory efficiency with text mining coupled with machine leaning. Besides discussing AI based NLP of literature data for safety assessment in Japan, Dr. Akihiko Hirose from National Institute of Health Science (NIHS) of Japan also discussed how the advanced AI methodologies improves some traditional techniques routinely used for risk assessment such as quantitative structure-activity relationships (QSARs), read-across and toxicogenomics. He provided a comprehensive review of various regulatory-driven applications of AI including development of chemical safety big database and AI platform to support human safety

assessment of pharmaceuticals, foods and household chemicals, deep learning model for Ames mutagenicity and hepatotoxicity.

Another question has been touched on by several speakers is the role of AI in public health, ranging from clinical diagnosis and prognosis, drug and food safety, disease prevention, precision medicine and nutrition. As we all aware that the COVID19 crisis has served as a boost for adoption and the increase use of AI in health care, what are some concerns/challenges and how to overcome the challenges. Drs. May Wang and Li Tong from Georgia Tech and Emery University of USA discussed five opportunities and challenges in AI for public health; these are (1) data integrity, (2) data integration, (3) causal inference, (4) real-time decision-making, and (5) metric and validation for explainable AI. Dr. Peng Li discussed regulatory science of medical device and AI in China with an example of applying AI-based image analysis for lung cancer and precision medicine. Dr. Kautz (NSF) described nEmesis, a foodborne illness surveillance system using a semi supervised leaning approach for analysis of Twitter data (Machine Learning epidemiology: real-time detection of foodborne illness at scale, Nature Digital Medicines, 1, 26, 2018) and the other AI system to extract information from social media to predict low self-esteem. Dr. Yinyin Yuan from Institute of Cancer Research of United Kingdom discussed AI for digital pathology in clinical diagnosis and prognosis. She touched on an important topic about reproducible AI for digital pathology. It was acknowledged that reproducibility takes multiple means, depending on the area of study, but in the context of digital pathology, controlling key parameters and measures are required towards improved reproducibility.

Weida Tong, PhD, National Center for Toxicological Research, FDA, USA

I propose a framework to consider for effective and efficient use of AI and machine learning technologies. The acronym of the framework is called **TRIAL,** which stands for **T**ransparency, **R**eproducibility, **I**nterpretability, **A**pplicability and **L**iability. Distinct from other similar frameworks from academia or other government agencies, the TRIAL framework is intended to serve as a reference point for AI applications that could be adopted across global regulatory community.

What are the elements of TRIAL and their inter-connection and complimentary features? Specifically,

- Transparency: An appropriate level of clarity for explaining algorithmic and data output of life-cycle AI performance allows assessment of AI by end-users and trust in the technology.
- Reproducibility: A set of well-defined measures ensure the trustworthy of an AI model.
- Interpretability: An AI model not only can be explained in human terms but also the causality of the underlying driving parameters to the model performance can be established with scientific support.
- Applicability: The context of use and application domain of an AI model need to be established including defining best practices, and whether AI can complement or replace other technologies.
- Liability: To prevent misuse of AI, the ethic rule and policy should be established which define the boundary of application and responsibility while promoting AI innovation in FDA.

As examples to articulate how the TRIAL framework would facilitate and guide the process of AI deployment in the regulatory setting, I provide the following steps: 1) State the problem that the AI algorithms will address, 2) Describe the data to collect in order to develop and validate AI results, 3) Replicate the results using different data sets if possible, 4) Design an AI platform for use in FDA or industry and to test feasibility in a pilot study, and 5) Implement the AI algorithm system-wide, monitor progress and update algorithm based on new data and results.

Stefan Platz, PhD, AstraZeneca, United Kingdom.

Integrating artificial intelligence (AI) in drug safety has the potential to unlock many benefits to enable us to enhance our capability, increase automation and accelerate our delivery.

I summarize my vision for AI as: "*AI will not replace toxicologists, but those who don't use AI, will be replaced by those who do*". This is an important statement because there's anxiety we will replace all toxicologists' work with a computer, which is far from the truth. But the toxicologist's role needs to evolve so people understand data more and use AI to find answers or collect insight that makes a more significant impact on our discovery and development.

There are three areas that have reached a level of maturity that enables us to maximize the momentum of AI:

- Computational power – with more powerful central processing units (CPUs) and Graphics processing units (GPUs), we can now manage and analyze huge amounts of data, which is key if we are to benefit from deep learning.
- Data connectivity – maximizing cloud-based systems, we can now access a wealth of data from different sources and from any location.
- Algorithmic development – combining intelligent algorithms with computational power and data connectivity, we can now access new insights into complex biology using multi-dimensional datasets.

Within AI there are three key areas: robotics, natural language processing (NLP) and machine learning. The latter has been available for a long time and includes linear regression, random forests or Bayesian approaches, so therefore, it's part of AI.

Many people have now been discussing deep learning, a subset of machine learning, where we take a different approach through multi-layered neural networks, which require a vast amount of data (Figure1).

How do we differentiate machine learning with deep learning? With the former, you define features (for example, to look for a triangle or a circle, etc.), set up rules and develop decision trees. You then cluster the information to reach an outcome. Despite the benefits with this approach, there are downsides, such as the system will only look for specific features you have defined, and it will ignore all the underlying information behind the data.

Deep learning takes this approach one step further and is very similar to how the human brain functions. Instead of feeding the system with 'features' (for instance, a triangle or circle), we build and train an artificial neural network with many thousands of reference images. The system learns to 'discriminate' by exploring different layers of analysis and setting up rules from all the raw image data. The key benefits are that you don't need to specify the features upfront, and you use all the information encoded in the data, which can be extrapolated with transfer learning. Essentially, you're going deeper and deeper in every layer, which is known as the multi-layered approach, and by connecting all the data, the computer delivers an outcome.

One of the hot topics in deep learning is the causality of prediction models – what is the relationship between cause and effect? Are the predictions of our models truly based on biological data?

We now have access to a hugely significant amount of data – internally and externally – but it's important to balance quantity and quality; having a huge volume of data doesn't always equate to relevant, high-quality information. It's important to structure your data and define ontologies, as you may have to merge internal and external data, and take extracts from multiple sources, such as medical data, product data, genomics data, imaging data and *in vivo* data. You can then consolidate data sources, some with their defined ontology and semantics, into a next generation sequencing platform.

Once you have structured your data, you need to connect the information into data mechanisms, such as data marts, knowledge graphs, data warehouses or data lakes. And from that point, you use AI technology in these data applications extract information. We use the acronym, **FAIR**, to describe the quality and usability of our data: **F**ind, **A**ccessible, **I**nteroperable and **R**eusable. The quality of your AI algorithms heavily depends on the data you feed in and how structured the data are.

1. Targeted safety assessment
The first example is in the **targeted safety assessment**. We find, in the phase two clinical trials, that we have attrition - fairly high on safety, but also in efficacy. There is an opportunity for AI to help increase understanding on a target you have picked to be LI (in the lead identification space) very early in the discovery, and to determine the connections of that target with other targets.

Within target safety assessment, we aim to develop a type of database called knowledge graphs which takes information from chemistry, pharmacology, clinical trial, genomics, multi-omics and biomarkers. In its simplest form, the knowledge graph connects two endpoints – a drug with a target – for example, gene X translates into protein X, and the graph provides some weighting to the strength of that connection. Extrapolating this can create a huge network which provides you with key information about connections through other targets and through other pathways. This may allow you to predict which target and off-target effects to expect.

2. Lead optimization in chem tox
Another example is the **lead optimization in chemical toxicity**. The question is, how rapidly can we optimize the structure and progress it to a point where we can test it as a medicine in humans? In tox studies, can AI help to make predictions, for example, a chronic tox study out of acute tox studies? Target organ toxicities? Pathologists? Imaging?

We have millions of molecules and the challenge has been selecting the right candidate quickly so it can be tested in humans. The influence and power of AI can help to show how you select other leads series that are similar in structure to molecules you know. Within AstraZeneca (AZ), we have made good progress in some of our predictions. We're developing an *in vivo* PK model where we can make predictions. It helps to get early data and narrow down the number of compounds that will be tested in *in vivo* PK data, which increases confidence in the selection of the right candidate.

Incorporating chem tox data can simplify matters for the chemist and toxicologist, who still have to make the decision. But visualizing all the different aspects on DDI and PK physchem property, general safety data, hepatic panel, cardiovascular panel, secondary pharmacology information and structure alerts – which is all available – can help people to understand the large volume of data to reach a better outcome.

Increased digitization and AI can help with the data we collect and share from our *in vivo* experiments. We are moving towards digitization, active monitoring, physiological end times and biomarkers. Active monitoring provides a more detailed picture and understanding on the *in vivo* experiments. Having access to more precise information can lead to more usable and translatable data to patients in clinical trials.

Digitization is a big step forward to increase the quality and output of our *in vivo* experiments, which really adds a translational aspect to it. This also reinforces our approach to identifying the right target, the right molecule, the right safety, right tissue and the right patient (Figure 2).

There is a substantial opportunity for AI in digital histology. We can simplify and accelerate the process by training the system, running a classifier and then checking the classified regions. With this approach, we have saved 90 percent of a pathologist's time by using AI.

We can look to the future and go one step further... if you take a tumor and then apply a mass cytometry (CyTOF) where we can look at 37 different markers, how do you analyze it? In this example, we have harnessed deep learning – we have created a neural network using software called t-SNE, which can indicate where there are more tumor cells. It connects and develops a complete description of how all the 37 different markers are improving our understanding on areas of proliferation and areas of different type of immune response.

If you then add another layer, which we did using mass spec imaging, you can even create a three-dimensional image of the tumor. There are different technologies being used from CyTOF to mass spec imaging, where a three-dimensional structure is providing you with detailed insights about active areas, hotspots, necrotic areas and inflammatory areas.

You may now understand that there is a greater need for toxicologists, but their role is evolving. They not only need to understand a wider range of modalities, but they need to realize the value and power of AI, and collaborate with experts to consolidate data, provide analytics and deliver insights to reach the correct conclusions and make the right decisions.

I am really excited about the opportunity that AI brings to discovery and development. It enables the delivery of meaningful insights that help us to keep pushing the boundaries of science to deliver life-changing medicines for patients.

Robert Ball, MD, MPH, ScM, Deputy Director of the Office of Surveillance & Epidemiology, Center for Drug Evaluation and Research, FDA, USA

In response to requirements in the Food and Drug Administration (FDA) Amendments Act (FDAAA) of 2007, the FDA launched an initiative to create the Sentinel System, a national electronic system for medical product safety surveillance. As shown in Figure 3, the Sentinel System became fully operational in 2016 [9] and continues to evolve to respond to FDA's needs [10]. The Sentinel system complements existing FDA surveillance capabilities, such as the FDA Adverse Event Reporting System (FAERS), that track adverse events reported after the use of FDA regulated products, by allowing the FDA to proactively assess the safety of these products. The Sentinel System includes the Active Post market Risk Identification and Analysis (ARIA) system. The ARIA system has two components: 1) a suite of preprogrammed, parameterizable and reusable analytic tools to enable rapid analysis of safety questions, including the use of sophisticated pharmacoepidemiological methods; and 2) data from the Sentinel common data model which consist primarily of quality-checked, electronic healthcare claims data (Figure 3).

Under the 2007 FDAAA, FDA could only require that a company conduct a post market safety study if FDA first determined that FAERS and the ARIA system were not enough to address the safety question. For the ARIA system, FDA defined sufficiency to mean that there is adequate data to identify: 1) the drug or biologic under study and an appropriate comparator when needed; 2) the covariates needed to address confounding; and 3) the health outcome of interest (HOI). In addition, the available parameterizable tools must implement appropriate

pharmacoepidemiological methods (e.g., propensity score-based confounding control) to answer the question of interest, to a satisfactory level of precision *(https://www.fda.gov/media/131980/download)*.

An analysis of the first several years of ARIA sufficiency assessments concluded that ARIA is sufficient for about 50% of the post-market drug safety questions for which CDER needed additional information [11]. The largest single reason for insufficiency is the inability to identify the HOI because many HOI require detailed data elements that are not available in the Sentinel CDM or because the algorithm used to create a computable phenotype does not have adequate performance. To address this limitation of the ARIA system, FDA began investigating artificial intelligence approaches, such as natural language processing (NLP) and machine learning (ML), to improve HOI identification as well developing a 5-year strategic plan (https://www.fda.gov/media/120333/download) for the Sentinel System that FDA has begun implementing.

As FDA encountered the insufficiency of HOI, two questions were raised: 1) whether incorporating key clinical features extracted from free-text fields in EHRs using NLP, combined with claims data would lead to better algorithms? and 2) Could we improve on algorithms constructed by human experts from claims data in the Sentinel CDM using ML?

To evaluate the first question, we selected anaphylaxis as a serious allergic reaction that occurs rarely after use of individual drugs but is a known adverse reaction across many drug classes. It is also a complicated diagnosis to make with certainty from EHR data because the diagnostic criteria require multi-organ system involvement, early treatment often occurs before full expression of the reaction, and diagnosis involves interpreting clinical information that is typically found in the narrative of the medical record. In a pilot project, we applied a previously developed NLP-based algorithm to cases of anaphylaxis identified in a prior Sentinel system study [12] (*FDA Sentinel System Innovation Center Master Plan. 2020. https://www.sentinelinitiative.org/news-events/publications-presentations/innovation-center-ic-master-plan).*
We found that this algorithm achieved similar performance to claims-based algorithms constructed by human experts. In conducting a qualitative error analysis of those findings, we observed that the algorithm was unable to make the same clinical judgment as human experts about timing, severity or existence of alternative explanations. To address this issue, Sentinel launched a comprehensive effort to develop a framework to use NLP and ML techniques to improve HOI identification algorithms, using anaphylaxis as an example. The results of this

project indicate that using NLP and ML it is possible to develop an improved algorithm to identify anaphylaxis in EHRs *(https://www.sentinelinitiative.org/news-events/meetings-workshops-trainings/electronic-health-records-natural-language-processing).*

To address the second question, Sentinel launched a project to demonstrate the feasibility and efficiency of developing and validating a claims-based HOI algorithm using machine learning classification techniques applied to a linked claims-EHR database *(https://www.sentinelinitiative.org/methods-data-tools/methods/machine-learning-pilot-electronic-phenotyping-health-outcomes-interest).* Rhabdomyolysis, a serious muscle injury that is a rare adverse effect of some drugs, was selected in part because it was possible to identify cases with a high degree of accuracy using laboratory data from the linked claims-EHR data, while having available the claims data necessary as input into the ML algorithms. This project successfully demonstrated that ML can improve the creation of computable phenotypes, even when features are restricted to claims data in the Sentinel CDM and do not include features constructed from the EHR.

The Sentinel system five-year strategic plan released in January of 2019 had several key messages. First, FDA aimed to maintain and enhance the foundation of the Sentinel system preserving FDA's long-term investment in the analysis tools and data infrastructure that form the foundation of the ARIA system. Second, FDA acknowledged the need to diversify and enrich the Sentinel data sources, especially EHRs and claims data linked to EHRs. Third, FDA proposed incorporating advanced analytics, including artificial intelligence approaches. Fourth, FDA looked to broaden the touch points for participating in Sentinel development to create a broader Sentinel community and disseminate knowledge generated by Sentinel to improve public health, ultimately fulfilling the goal of making Sentinel a national resource for multiple purposes.

While the examples provided earlier in the manuscript focus on NLP and ML for improving HOI identification, Sentinel System has proposed a comprehensive set of innovation strategies (figure 2) to address the full range of data, methods, and tools required to build a sustainable infrastructure that incorporates artificial intelligence methods to address FDA's needs. In October of 2019, FDA launched a new contract to implement the vision of the five-year strategic plan that included three Sentinel centers: an operations center, an innovation center and a community building an outreach center. The Sentinel innovation center has created a 4-pronged approach (Figure 4) to implement the Sentinel System innovation strategies, including NLP and ML for constructing computable phenotypes (https://www.sentinelinitiative.org/news-events/publications-presentations/innovation-center-ic-master-plan).

The use of Sentinel to respond to the COVID-19 pandemic has highlighted the need for improved and rapid access to EHR and linked claims-EHR data to address questions about drug treatment in hospital and especially intensive care unit settings (*FDA Sentinel System's Coronavirus (COVID-19) Activities. 2020. https://www.sentinelinitiative.org/assessments/coronavirus-covid-19).* The need to respond to this public health emergency is further informing the implementation of the Sentinel strategic plan.

In summary, FDA has been evaluating artificial intelligence technologies especially natural language processing and machine learning, combined with improved access to detailed clinical data from EHRs to improve post market drug safety surveillance in Sentinel. The FDA five-year Sentinel System strategic plan outlines this vision and is currently being implemented Figure 5).

Philippe Girard, PhD, Deputy Executive Director of Swissmedic, Switzerland

Swissmedic is moving forward in the field of artificial intelligence starting with a pilot to detect safety signals using artificial intelligence in the context of clinical trials.

Swissmedic approves around 200 clinical trials every year. Approximately 1 million biomedical scientific publications focusing on clinical trials and their effects and adverse events are published yearly. Lack of resources to monitor these publications is the reason why the discovery of relationships between published adverse events and our studies has been random luck if not reported by the sponsors. Obviously, we are aware of the safety benefits for participants in clinical trials if we could reliably connect the trials in Switzerland with the publicly available safety data. The solution could lie in using new technologies such as natural language processing, artificial intelligence (AI) and machine learning to allow for a better safety monitoring.

Swissmedic decided to establish a proof-of-concept (PoC). For our retrospective study, data was gathered from seven open-source websites *(https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts/archive-recalls-market-withdrawals-safety-alerts; https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts; https://www.fda.gov/recalls-market-withdrawals safety/-safety-alerts; https://www.ansm.sante.fr/S-informer/Travaux-de-l-Agence-Europeenne-des-Medicaments-*

EMA-Comite-pour-l-evaluation-des-risques-en-matiere-de-pharmacovigilance-PRAC;
https://www.ansm.sante.fr/S-informer/Informations-de-securite-Lettres-aux-professionnels-de-sante; https://www.gov.uk/drug-safety-update; https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management/prac-recommendations-safety-signals)

and five investigational medical products (IMP) of concluded clinical trials were considered. We developed a set of filters to narrow down the list of articles for analysis. These are: Filter 1: Is this article drug-related?
Filter 2: Is this article related to the disease indication we are looking for? Filter 3: Is this article related to drug related adverse events (vs only symptoms of the disease)? Filter 4: Is this article related to safety signal (Swissmedic definition)? Filter 5: How relevant/ seriousness level (High/Medium/Low) is the detected sentence? (*https://www.ncbi.nlm.nih.gov/pubmed*) To validate the approach, the results of the AI system were compared against those of a human reviewer.

The PoC consisted in three steps. First, cloud tags were to be defined. To do so, synonyms of the respective IMP and adverse events were compiled. These lists were generated and cured by a human reviewer. In total, the cloud tags contained more than 14,000 different terms. With these cloud tags, the publication websites were crawled for the relevant keywords. In the resulting texts, the position of the cloud tags was determined relative to each other to generate structured data the engine could use for the further steps. The results of this step were ranked, and a sample was iteratively plausibility-checked by a reviewer who had to ensure that the identified adverse events were related to the IMP mentioned in the publication. By using filters, like 'is the article drug-related' or 'does the indication match the clinical study', the results were further cleaned in order to remove unrelated publications. Finally, the importance of the publications was ranked (Figure 6).

The results of the IMP safety monitoring PoC were evaluated in three categories:
Comprehensiveness: Within the PoC, the AI identified all 10 hits that were also identified by the reviewer. Moreover, the AI identified 14 hits that were not identified by the reviewer. The reviewer was subsequently unable to find one of those AI-detected hits through directed manual search.

Relevance: All hits identified by AI were relevant. On the other hand, only 2% of the initially identified manual search hits by the reviewer were relevant after analysis.
Benefit: The reviewer invested 145 min for manual search and initial evaluation versus only 10

minutes for the AI. We see also a benefit of the comprehensive in-depth document analysis of the publication generating more trust in the results as compared to a limited manual review.

In summary we can say that AI methods can efficiently support monitoring and the detection of adverse events in clinical trials. Moreover, the results were deemed reliable as well as relevant, and therefore highlight the potential time savings within the review process.

Intending to explore the possibilities of the digital transformation in-depth and with a high priority, Swissmedic has launched the digital initiative Swissmedic 4.0 in June 2020. It is an agile and independent unit equipped with sufficient resources to be set up for this purpose. The digital initiative will ensure that Swissmedic has the resources and facilities to tackle the challenges of digital transformation regardless of the limitations imposed by daily business. Questions arising from the overall organization will be received, analyzed, and processed, and the resulting solutions will be fed back into the organization. The initiative will learn from its mistakes and improve its proposed solutions. This will allow Swissmedic to actively tackle the great changes of the digital transformation as an opportunity and not just passively allow them to pass by. The focus will not be limited to the technical aspects. Swissmedic's digital transformation challenges should rather be anticipated with a comprehensive and sustainable approach.

This experience is directly incorporated into further AI implementation currently under investigation by the Swissmedic 4.0 team, such as automated request-response. With an automated Q&A tool for classifying emails and the associated conception of a chat offer, Swissmedic hopes to gain valuable knowledge and experience regarding neurolinguistics programming (NLP) methods on the one hand and to relieve our communications department by filtering and channeling audience enquiries on the other.
Swissmedic is exchanging, and we plan to further intensify these exchanges, with health authorities, stakeholders, and the science communities in order to accelerate these developments.

Blanka Halamoda-Kenzaoui, PhD, Joint Research Centre, EC

A feasibility study on the use of automation tools for a systematic review of the scientific literature is presented. In the field of innovative health products such as nanomedicines, high quality published data could provide the lacking information on still open questions, supporting

the release of regulatory guidance on data requirements and decreasing the uncertainty for product developers. However, monitoring of the scientific literature, which is constantly increasing, and processing of high amount of data requires the use of automation that can speed up the process and reduce a bias.

In order to identify specific toxicity effects induced by nanotechnology-enabled products, a battery of available automation tools for initial steps of the systematic review process such as targeted searches, de-duplication and screening of abstracts is used (Figure 7). For more specific steps which are related to the evaluation of article quality and extraction of information no ready-to-use tools are available. Our in-house developed segmentation tool allows extraction of specific sections of scientific articles before applying further text mining techniques. A set of criteria related to reported nanomaterial characterisation is applied to the Materials and Methods section in order to score and rank the articles. Furthermore, based on the developed ontology, the most reported toxicity effects are extracted and mapped against different types of nanomaterials.

Such knowledge is of critical importance for the development and regulatory assessment of nanotechnology-enabled products. The study confirmed the huge potential of automation systems especially in the area of innovative health products, however, further improvement of their reliability and applicability is needed (Figure 7).

Yinyin Yuan, PhD Computation Pathology and Integrated Genomics, Centre for Evolution and Cancer, Institute of Cancer Research, United Kingdom.

By examining cancer cells in pathological samples, one can begin to understand the tumor landscape including the diverse coexisting ecosystems. In these samples, cancer cells can be seen to be surrounded by immune cells and vessels in some parts of the tumor, and in another part, cancer cells just seem to be living on their own. Using AI, we can automate analysis of these images to map out these different tumor components and analyze them in detail.

In search for reproducible AI for the studying lung cancer in the program TRACERx [13], we developed AI tools to study immune response using digital pathology data [14]. TRACERx is a prospective study of 842 non-small cell, lung cancer patients, tracing them from diagnosis to relapse. The aim is to understand why some patients develop drug resistance and cancer relapse.

By studying the interaction between cancer cells and immune cells in the tumor, immunotherapy

was developed, and subsequently, the checkpoint-based therapy has changed the cancer treatment landscape. They induce prolonged tumor response through reactivating immune cells. However, most patients still don't benefit from this immunotherapy.

On the other hand, targeted therapies exploit the genetic vulnerability of cancers, such as kinase inhibitor targeting in the mutated cells. However, just like a virus, cancer cells can evolve and acquire genetic variation within a patient. This means that cells or subclones that evolve resistance, are selected, diversify, leading to metastasis and cancer recurrence. The question is, "How can we better understand and treat cancer by studying immune response and genetic heterogeneity?" and importantly, "How to translate this new knowledge in the benefit of the patients?"

The TRACERx team has performed DNA sequencing on multiple regions sampled on surgically resected lung tumors. By comparing the genetic profiles of these regions within the same tumor, the genetic study revealed a fundamental biological feature of lung cancer, which is the widespread genetic intratumor heterogeneity. A phylogenetic tree of cancer can illustrate how cells acquire early and late clonal mutations and diversify genetically in one patient. The question remains, "What promotes such genetic heterogeneity and what is the role of immune cells in cancer evolution towards genetic heterogeneity?

As with any AI project, data quality is key to success. We developed a specific protocol to optimize data quality for both pathology and genomics data. The multiregional samples collected for TRACERx are punched biopsies from surgically resected tumors, and then frozen. While this is good for DNA sequencing, tumor morphology is not well preserved in frozen sections and often fragmented. Therefore, half of the frozen tumor block were re-embedded in paraffin to generate good quality pathology data.

We then applied deep learning to pathology images to create an AI tool capable of distinguishing immune cells from cancer cells. This tool allows to map the spatial distribution of cancer cells, lymphocytes, and other cells. To make sure that the tumor doesn't overfit the data, the tool was trained on a diversity of sample types generated in TRACERx to capture as much variability as possible and tested extensively on external cohorts.

The difficulty, then, is to get sufficient ground truth data. Our model was trained by human intelligence that can translate years of pathological experience. How can we deliver rapid validation with the independent samples and objective methods? We developed a new way to

generate quantitative data and scale with biological validation. In this method, we first stained samples with immunohistochemistry using cell type as a second marker and then stained with clinical routine pathological, hematoxylin & eosin (H&E)-stained samples. One experiment can generate enough data that would have taken hundreds of hours for a pathologist to annotate. With this approach, we can directly relate cell types identified by protein expression with H&E automated cell classification to show that there is a good concordance between the two.

What we found is, regardless of how many regions were sampled, the number of immune cold regions that came out was the most predictive of relapse in lung cancer in all the immune features tested. This was first discovered in TRACERx, and later validated by applying the same AI model on the largely suspected cohort of nearly1,000 adenocarcinomas consisting of 4,000 cold tumor slide images.

In both cohorts, the increased number of immune cold regions is associated with a high risk of cancer relapse, even for a tumor that has an average of thousands of lymphocytes. It's also independent of all the clinical parameters including smoking history and age. This means that the stage and location of immune cells is more important than the numbers, but why?

By integrating images with genomics data, this study created a detailed picture of how lung cancers evolve in response to the immune system in different patients. Cancer cells in immune hot regions tend to diversify into different genetic subsets early in their evolutionary history, perhaps under intensive selective pressure. Whereas cancer cells in the immune cold regions tend to diversify later, perhaps due to the ability of immune escape.

While this remains to be validated in our next stage of the study, it provides evidence that the AI tool for repeat verification of lung cancer is reproducible and might have uncovered a key biological basis. It also means that, by focusing on the genetic subsets where subclones in immune cold regions start to diversify, we may be able to find new ways to target immune -- cancer escape from the immune surveillance. The hope is that this AI tool could be used in the future to pick out lung cancer patients at highest risk of relapse, helping inform a more tailored treatment strategy.

AI also has a utility for early detection of cancer. In a recent study [15], we showed that the AI immune scoring method trained in lung cancer, can be used to predict precancerous lesions in the lung. About half of our patients with carcinoma in situ lesions in the lung progressed to have lung cancer. The question is, "How can we better predict who will progress for personalized

medicine"? By using our AI tool to score immune cells, we found that carcinoma in situ lesions that later progress have less infiltrating immune cells compared with those that later regressed. This AI tool may even help stratify these patients to help identify who may need intensive monitoring, and the lack of immune cells for these immune cold phenotypes may also be relevant in the evolution of precancer to cancer. To summarize, AI not only has a large range of applications for pathology and medicine, but it also generates new demands by enabling what was impossible before.

Akihiko Hirose, PhD, National Institute of Health Sciences), Japan

The focus of this research is the development of a risk assessment support data system. The goal is to support risk assessment of chemicals, food, or medical ingredient and additives. And the system output may assist risk assessor.

We have developed several toxicity information databases for supporting each regulatory field. We thought that an AI-base toxicity evaluation system could be established by integrating these inhouse databases. At first, we introduce three original databases in this workshop. The first one we named the HESS database. This tool can support read-across assessment of chemicals like OECD tool books and is specifically focused on repeated dose toxicity. The second is the Ames test database containing more than 10,000 chemical research. The third is the toxicogenomic database. We have developed the standardization method to capture gene expression data called the Percellome method. Now we are trying to integrate the database into the common research platform as a tool for open toxicology. Finally, I will introduce the concept and the plan of the A.I.-based Chemical Safety Assessment Forward Evaluation platform and present activity of development of pilot prediction system as a prototype system.

The first topic is repeated toxicity database. We have developed a hazard evaluation support system, HESS platform. HESS is based on the database of repeated dose toxicity study results but includes a knowledge base of toxicology mechanism, metabolic maps, and metabolite simulators in rats, and ADME database in humans and rats. The system can show the information from the database and the toxicological category and predict the toxicological parameters. The series of outputs can support expert judgment. We developed the read-across case study at OECD IATA in 2018 by using this HESS platform. The case study focused on the testicular toxicity of ethylene glycol methyl ether and related chemicals with a common active

metabolite.

In addition to the HESS system development, we are developing another type of database for improved performance of toxicity prediction. To achieve this goal, we are now gathering the test data of new chemicals, food additives, food contact materials, pesticides, pharmaceutical additives, and so on. As for developmental and reproductive toxicity assessments, we used the OECD 421 and 422 study test data under the Japanese existing chemical program. We also searched possible molecular targets and develop AOPs by using this database.

The second topic is Ames test database. We have accessed to Ames test data for many chemicals collected over the last 30 years under the Industrial Safety and Health Act in Japan. This Ames NIHS database includes 12,000 chemicals.

Using this large data set, an international collaborative study was proposed in 2014 for improving the Ames QSAR models. This project has three trial cycles, Phase I to III, and includes about 4,000 substances in each trial challenge. Twelve models were used in this project. The specificity and sensitivity generally increased with increasing phase and saturated at Phase III. One reason for this may be the variability in biological testing. Some reports say that reproducibility of Ames test is 80 or 85 percent. We think that it is important to use reliable training data for improvement of the models, and to carefully evaluate the mechanism of toxicity. After the Ames test dataset was revised and QSAR models improved, the QSAR is no longer just a prediction system but a genetic toxicity assessment to system beyond the Ames test.

As for the third topic, we developed a toxicogenomic database with a normalization method to generate absolute copy number within mRNAs. The Percellome method we developed is a spike in method. We measured the number of the cells in the sample homogenate and spiked in five kinds of bacterial mRNA according to the cell number. Each homogenate included mRNA copy numbers covering a range of about one to 100 copies per cells. This method enables direct comparison expression data between various experiments data. Forty-eight samples are measured by 48 gene chips and normalized. 45,000 kinds of mRNA can be loaded in.

We are now developing the prediction method of repeated toxicity from the existing data of the Percellome single-dose gene expression analysis. In addition to the gene expression data, we also can analyze epigenetic data. Additionally, we have integrated the rat gene expression data to our mice database system. Moreover, we provided a web-based API for integration to the common platform of bioinformatic software.

By integrating these larger scale reliable toxicity data, our institute has started to develop a large chemical safety database and A.I. platform for supporting efficient and reliable human safety assessment for pharmaceuticals, foods, and household chemicals and for providing expert knowledge beyond regulatory frameworks.

At the first stage, we are constructing the prototype platform by using three kinds of databases: HESS database, Ames database, and public domain assessment document database. Our first project, the development study of deep learning models for predicting Ames mutagenicity, is being conducted as a pilot study because the Ames mutagenicity test data is the most abundant among the many toxicity test data. The prototype model automatically extracts the feature by using SMILES of chemical structure and convolutional neural network. The prediction accuracy was about 80 percent using this method. As alternative method, we are developing uses a new deep learning model using graph convolutional network.

As our next target, we are building a new prediction model for hepatoxicity by integrating various in silico and in vitro test data. First, deep reading using a toxicity study data set and only structural information. Second, deep learning using data flux from in silico in vitro data such as AOP event information or in vitro bioactivity or metabolite information, in addition to the structured information. A positive prediction was about 0.68%, sensitivity was 0.75, f1-score was 0.71. When the structured information and in silico in vitro data types were used, the model performance provides a relatively good score in learning.

Because literature searching is time and resource consuming, we developed a tool for text searching based on the practical application of highly accurate search technology for natural language text information and expert judgement. We used sparse composite document vectors and conducted a clustering and topic analysis by distance calculation in vector space. Our preliminary results indicated a correct answer hit 0.62.

In conclusion, the systems we are developing serve as an administrative measure to ensure strong risk management of pharmaceuticals, foods, and chemicals. It is essential to establish a new method for accurate and efficient safety assessment support system. Physical chemical and exposure data are also essential for the risk assessment. Furthermore, we believe it is necessary to improve the performance of the system by conducting verification experiments on the data of the generated by the prediction system.

Li Tong, PhD and May D. Wang, PhD, Universities of Georgia Tech and Emory, USA

Biomedical researchers are developing and advancing biomedical technologies over the last few decades towards predictive, personalized, participatory, and precision health, or abbreviated as pHealth (Figure 8). With the emergence of high-throughput health data, artificial intelligence (AI) has been utilized to make sense of the data. However, to translate AI models to clinical practice, it is critical to understand whether the model is reliable and reproducible.

From the regulatory perspective, it is also essential to ensure whether the results or the method

itself is interpretable. Only after stratifying these regulatory concerns, we can make a joint decision about any patient or cohort of the population by integrating multi-modal data with AI models. Unfortunately, compared with the fast development of AI-based models, establishing regulatory frameworks for these models remains a challenging task [16, 17].

With the accumulation of multi-omics data, computational biology and bioinformatics are playing essential roles in extracting meaningful information from the vast amount of raw -omics data. Hundreds of thousands of new bioinformatics tools are developed each year for various -omics data (e.g., RNA-seq and DNA methylation) and applications (e.g., gene expression estimation). For example, after extracting short reads from the biopsy samples using the next-generation sequencing (NGS) technique, various bioinformatics pipelines are utilized to extract features such as gene expression levels from these raw RNA-seq data (Figure 9A and 9B). However, very few of these bioinformatics tools have been translated to clinical practice. One major challenge is the evaluation and selection of numerous bioinformatics pipelines. For the application to clinical practice, bioinformatics tools should form into pipelines with proven effectiveness and reproducibility. The quality control and supervision led by authorities such as the United States Food and Drug Administration (FDA) play an irreplaceable role in this process. These agencies are supposed to provide broad regulatory oversight and validation for bioinformatics tools applied to clinical practice by forming evaluation standards and supervising clinical validation.

RNA-seq is one of the next-generation sequencing technologies widely applied in biomedical research. However, the translation of RNA-seq for medical and clinical applications requires selecting reliable and reproducible bioinformatics pipelines. The US FDA has led the Sequencing Quality Control (SEQC) project to conduct a comprehensive investigation of 278 representative RNA-seq data analysis pipelines consisting of 13 sequence mapping, three quantification, and seven normalization methods [18]. As a follow-up study, we further

investigated the impact of the joint effects of RNA-seq pipelines on gene expression estimation and the downstream prediction of disease outcomes [19]. First, we developed and applied three metrics (i.e., accuracy, precision, and reliability) to evaluate each pipeline's performance on gene expression estimation quantitatively. We then investigated the correlation between the proposed metrics and the downstream prediction performance using two real-world cancer datasets (i.e., SEQC neuroblastoma dataset and the NIH/NCI TCGA lung adenocarcinoma dataset). We found that RNA-seq pipeline components jointly and significantly impacted the accuracy of gene expression estimation, and its impact was extended to the downstream prediction of these cancer outcomes. Specifically, RNA-seq pipelines that produced more accurate, precise, and reliable gene expression estimation performed better in predicting disease outcomes. Thus, the evaluation and selection of proper bioinformatics pipelines are important to ensure meaningful downstream analysis such as the predictive modeling.

Besides evaluating bioinformatics pipelines' effectiveness and reproducibility, another component requires government agency-dominated quality control and supervision lies in predictive modeling (Figure 9C). AI and machine learning have achieved unprecedented success in various applications, including computer vision and natural language processing. Although these AI techniques, especially deep learning, have been widely adopted in biomedical research [20, 21], clinical practice penetration is still limited. The translation of AI-based predictive models faces similar but more challenging regulatory issues compared with that of bioinformatics pipelines.

The first quality control issue of these AI-based predictive models is the model reliability and reproducibility. Hundreds of thousands of new models are reported every year, with each of them claims state-of-the-art performance. However, most of the reported performance is probably too optimistic. To translate these models to clinical practice, we need to establish broad regulatory oversight and validation, collaborated by both authorized agencies and research communities. For example, the risk of bias is a major issue for AI-based predictive models. One potential

direction is establishing a standard evaluation dataset and metrics containing datasets from various sources and populations to minimize the risk of bias. These benchmark data should be aggregated and shared among entities by agencies such as the FDA. On the other hand, the medical AI engineers and researchers should also utilize meta-learning and transfer learning techniques to improve the model's generalizability from the computational side.

The second issue of predictive models that requires supervision is model transparency. It is essential to build trust among clinicians and patients. Instead of a black box, we need to understand how the model makes decisions, why it works, and when it could fail. One major effort to address this issue is the development of interpretable AI (XAI). XAI is introduced to ensure humans can fully understand the model and every logic behind the final decision can be inspected, audited, and trusted [22]. From the regulatory perspective, the supervision agencies should provide guidelines and frameworks for engineers to enable XAI for each step of model design, implementation, training, and deployment.

The third issue that needs to be supervised is the privacy risk. The privacy constraint is one of the major challenges for biomedical data sharing, cloud computing, and model deployment. Researchers have developed various privacy-preserving deep learning techniques (e.g., federated machine learning) to solve this issue [23]. However, the regulatory authorities must play an essential role by establishing guidelines and model interoperability resources.

AI-based predictive modeling is the key to integrate multi-modal biomedical data and further enable pHealth. However, several gaps (i.e., model reliability and reproducibility, model transparency, and privacy risks) need to be addressed before these novel techniques can be confidently translated to clinical practice. The regulatory agencies can play a key role in this process by collaborating with academic institutions and commercial companies to build a sustainable ecology for the application of AI in healthcare.

Track B: Omics, Biomarkers, and Precision Medicine in Regulatory Science: (Neil Vary and Primal Silva, PhD, Canadian Food Inspection Agency, Canada and Richard Beger, PhD, National Center for Toxicological Research/FDA, USA and Susan Sumner, PhD, University of North Carolina, UNC Chapel Hill, USA)

Molecular biology specializations in "omics" (genomics, proteomics, and metabolomics) are continually advancing as various private, academic, and government organizations invest to support research in these areas. The development of novel technologies, applications, and methodologies are being incorporated into regulatory science laboratories, such as food production and precision medicine, as these scientific fields continue to evolve. These specializations are being used to identify specific biomarkers of interest, such as antibiotic resistance in foodborne pathogens, and to identify targeted genetic sequences, protein biomarkers, and metabolites used to advance precision medicine.

The adoption of genomics analyses of pathogens in microbiological laboratories, including regulatory laboratories, is becoming common. Genomics is now being used to confirm the identify of pathogens, characterize them and match genetic sequences of clinical isolates to foodborne isolates, and identify specific genetic sequences of interest (e.g., toxin genes, antimicrobial resistance [AMR] genes, and serotypes). Genomics is also used to detect genetic disorders which can lead to the development of precision medicines, for both rare diseases (where there may be no therapies) or for stratification of common diseases. Pharmacogenomics is a specific area of research where genetic variation in genes determining the pharmacokinetics and/or pharmacodynamic of drugs can lead to variation in drug response, both efficacy and safety. There is an increasing global trend towards implementation of pharmacogenomic variation into clinical practice, with guidelines being produced by organizations such as the Clinical Pharmacogenetics Implementation Consortium.

Like genomics, proteomics approaches are beginning to be explored and integrated in regulatory laboratories. Proteomics can be used for food traceability and quality, as well as food safety to screen for foodborne pathogens or allergens with high sensitivity and specificity. Research in proteomics is also used to characterize protein structures of viruses, such as SARS-CoV-2, that may contribute to the development of vaccines.

Currently, metabolomics is used to diagnose complex metabolic diseases, and the field is growing rapidly. Metabolomics is being used to help further efforts to develop and advance precision medicine applications.

This track will focus on the evolving "omics" fields and how they are being used by regulatory bodies through incorporation into regulatory science laboratories, and how they can be further exploited as these fields continue to mature.

The first half of Track B focuses on genomics, proteomics, and precision medicine.

Dr. Tim Mercer from the Garvan Institute in Australia described his team's efforts to measure and understand the human genome using next-generation sequencing technologies. This included the approaches that can be used to validate next-generation sequencing technologies for clinical use, and how they applied those same tools and concepts for a large evaluation of liquid biopsy assays that were completed recently as part of the SEQC2 consortium.

Dr. Munir Pirmohamed, a professor of medicine at the University of Liverpool and the NHS Chair of Pharmacogenomics delivered a presentation about pharmacogenomics and precision medicine. Precision medicine allows more effective drug therapies for patients, as it is recognized that individual patients do not respond equally to prescribed medicines. Some drug therapies are taken by patients for long periods of time but have no effect on their disease. In many cases, patients have adverse reactions that can cause serious harm and require further medical intervention. Precision medicine can break down general diseases into sub-types that respond differently to different medications. It can also be used to look at specific variables in individual patients, including genetic factors that impact drug effectiveness.

The second half of the Track B session is focused on metabolomics. Dr. Richard Beger describes a typical workflow for metabolomics and then discusses current public outreach in quality assurance (QA) and quality control (QC) in metabolomics. He also describes how metabolomics has been used to identify translational biomarkers of acetaminophen induced hepatotoxicity in mice, rats, and humans.

Dr. Susan Sumner, Department of Nutrition, University of North Carolina at Chapel Hill describes the use of a metabolomics/exposome platform to reveal significant metabolic perturbations that arise for individuals using opium compared with non-opium users. The investigation also pointed to biomarkers associated with opium use disorder. She postulated that a nutrient cocktail of vitamins or their metabolites, vitamin-like compounds (such as choline), and fatty acids may protect against metabolic disruptions that lead to addiction -consistent with the Doyle-Nyswander theory that addiction is initiated by a metabolic imbalance.

Dr. David Wishart, University of Alberta, Canada, shared his work on software and databases that have been developed to enable omics-based regulatory science, with a focus on metabolomics. The most widely used and best-known resource developed by his lab is called DrugBank. It's a drug database and currently has around 2700 small molecule drugs, 1400 biologics, 130 nutraceuticals, and includes more than 6300 experimental drugs. A food constituent database called FooDB has almost 80,000 compounds found in 730 different foods.

Dr Reza Salek, International Agency for Research on Cancer (IARC) in France, which is part of WHO/United Nations described how metabolomics and exposomics, particularly cancer epidemiological is appledat IARC. Additionally, they use -omics data integration approaches from a pathway centric points of view for interpretation of the results.

Dr. Hairuo Wen from National Institute of Food and Drug Control, China, focuses, on using the NSG mouse model, nonclinical evaluation of chimeric antigen receptor modified T-cells against CD19. The study provides a comprehensive report of the efficacy, biodistribution, toxicity, and especially the immunotoxicity of CAR-T using an NSG and Raji-cell xenograft lymphoma model. The NSG mice that received CART19 treatment demonstrated a longer survival period without significant immunotoxicity, suggesting and encouraging clinical prospect of CART19.

TimMercer, PhD, Garvan Institute, Australia.

Next-generation sequencing (NGS) have become widely used in biomedical research and is being increasingly adopted for clinical diagnosis. Accordingly, there is an increasing need to understand the performance and limitations of NGS-based tests in clinical diagnosis.

The validation of an NGS-based diagnostic test requires understanding two major parameters. Firstly, sensitivity; the ability to detect true-positives, such as the detection of rare mutations, and secondly, specificity; the ability to not erroneously detect false-positives. In NGS, we typically use precision instead of specificity due to the large number of true negatives calls that result in unbalanced classes.

The relationship with sensitivity and precision is rarely linear, simple, or obvious. There is typically trade-off between sensitivity and precision (See Figure 10). The ideal tradeoff between these parameters depends on the aims and the goals of a diagnostic test. For example, some NGS tests may require high sensitivity, such as the diagnosis of rare somatic mutations, whilst other NGS tests may require high precision, such as whole-genome sequencing, where even a

small false-positive rate can result in many errors across the entire genome.

The sensitivity and the precision of an NGS test can be measured using reference standards, genetic materials with well-known ground-truth properties [24]. Natural genetic materials provide useful reference standards as they are typically inexpensive, and they typically encompass the full size and the diversity of the human genome or the transcriptome. However, they must be well characterized using orthologous genome technologies, and can be difficult to sustainably manufacture. Whilst the development of reference genomes, such as the NA12878 genome, have been key in evaluating whole-genome sequencing, the development of cancer reference standards has been more problematic [25].

The widespread availability of DNA synthesis has also enabled the development of synthetic RNA and DNA controls. These controls can be rapidly designed and flexibly manufactured to represent a specific genetic sequence and can be mixed at different concentrations to represent quantitative genetic features, such as varying allele frequencies or copy number variation [26]. Finally, different synthetic controls can be combined into a single mixture, enabling the incorporation of many different diagnostic features into one mixture that matches the breadth of features that can be diagnosed using NGS.

With these reference standards, how do we proceed with the clinical and the analytical validation of an NGS test? As an example, I will report the findings of a large multi-lab, cross-platform study that evaluated the performance of circulating tumor DNA assays using a range of reference materials.

Circulating tumor DNA (ctDNA) are DNA fragments that are released by cancer cells into the blood stream. CtDNA can harbor somatic mutations that indicate the tumor of origin, and their abundance can indicate tumor size and stage [27]. The collection of ctDNA is rapid, inexpensive and minimally invasive, and can be performed serially to monitor tumor evolution and therapy response. Given these advantages, there has been considerable attention and investment in using ctDNA as an accessible cancer biomarker.

Despite its potential, ctDNA assays faces some major technical challenges. The ctDNA fragments exist at low concentrations, there is a large amount of non-ctDNA circulating in the bloodstream and detecting rare somatic mutation from a limited amount of input material can be challenging. Given the clinical adoption of ctDNA assays, there is a pressing need to measure the sensitivity and the precision of these assays, and to understand the variables that impact the performance of ctDNA assays.

We initially simulated ctDNA NGS libraries that have undergone targeted capture and enrichment. The benefit of simulated libraries is they allow us to understand the basic

parameters that frame ctDNA analysis and can distinguish the impact of bioinformatic variables from downstream experimental variables. For example, we could use simulated libraries to evaluate the impact of where somatic mutation occur within an exon or understand the impact of read alignability in low-complexity genetic sequences.

We next employed synthetic DNA controls that represented known and important cancer mutations. These synthetic controls were diluted at decreasing concentrations to form quantitative scales of ladders of different allele frequencies. That enables us to measure the limit of detection, the sensitivity for rare allele frequency and quantitative accuracy of the assay. Notably, we found a limit of detection at a .05 percent allele frequency, with detection of mutations below this abundance being more variable and uncertain. This limit requires increasingly greater sequencing depth to improve and may represent an inherent limit of ctDNA assays.

We finally evaluated the reproducibility and the reliability of ctDNA assays across different technologies and different laboratories. To achieve this, we organized this large-scale proficiency study using a mock tumor sample that comprises different cancer cell lines mixed at different frequencies. This mock tumor sample was provided to participating laboratories who perform ctDNA sequencing and analysis according to their standard protocols, with the results sent to a centralized location for analysis and benchmarking of performance.

This proficiency study evaluated five different commercially available ctDNA assays, each which used a range of different technological methods or approaches, that was performed across 13 different laboratories. At completion, the proficiency study encompassed 360 different circulating DNA tests representing, to my knowledge, the largest proficiency test for ctDNA assays, and probably one of the larger proficiency tests for NGS.

Similar to our results from the synthetic control above, ctDNA assays rarely achieved a sensitivity to detect mutations below .05 percent allele frequency threshold. However, we did find that precision was uniformly high across the ctDNA assays, and the use of unique molecular identifiers resulted in very few false-positives mutations detected. However, there is a tradeoff to this high precision, with the sensitivity for indels was quite low, and it is likely that a higher sensitivity for indels will result in lower precision.

Nevertheless, most assays performed robustly and reliably across the laboratories, and we found that steps in the workflow, from the plasma extraction to the library preparation, to the sequencing and the bioinformatic analysis, were generally impacted by only a few random, rather than systematic variables between the laboratories. Therefore, I would encourage researchers establishing of developing ctDNA assays to take a closer look at the large amount of

data generated within this study, as well as the broader recommendations we made, when establishing and validating ctDNA assays within their laboratories (Figure 10).

Munir Pirmohamed, FRCP, PhD, University of Liverpool, United Kingdom

There is a huge degree of variability in drug efficacy – more than 90 percent of drugs only work in 30 to 50 percent of people. The trial-and-error approach we currently use means that for some patients it takes a prolonged period of time to identify the right drug that works for their disease. Conversely, there are also some patients who develop adverse drug reactions (ADRs): studies have shown that 6.5% of all admissions to hospitals are due to ADRs [28], and 15% of patients in hospital develop ADRs [29], all of which has an enormous cost burden on healthcare systems (in addition to causing morbidity and mortality).

When one examines the role of precision medicine in drug response, it can be divided into two broad areas. The first relates to disease classification – we still use the same classification that was developed in the 19th and 20th centuries and is based on phenotypic criteria. However, molecular techniques are beginning to show there is a great deal of heterogeneity in diseases, such that in the future there may be a need to change the taxonomy of disease. This also opens the possibility of disease stratification with different disease sub-phenotypes responding to different therapies. For example, in asthma, stratification into different disease phenotypes is leading to the development of precision medicine approaches [30]. Similarly, in cystic fibrosis, different mutations in the CTFR gene lead to differential disease phenotypes and the use of different combinations of precision therapies. The second relates to variability in drug response even in individuals in the same disease strata. This area of research is called pharmacogenomics, the study of variation in DNA and RNA and how this determines drug response.

Stratification of disease can also be based on the somatic genome, as has been shown with cancer. Cancer is a genetic disease, where the somatic genome deviates from the germline genome through the occurrence of mutations in many genes, including key driver genes. This has also been opportunistic as it has provided the ability to develop targeted therapies, for example vemurafenib for the V600E mutation in the BRAF gene in malignant melanoma. The development of targeted therapies is now routine in many different cancers [31], and is used together with conventional chemotherapeutic agents and immunotherapies to improve response rates in many different malignant conditions, sometimes with impressive results.

With the advances in genomic technologies, it is now easier to undertake whole genome sequencing, which provides information on all pharmacogenes [32]. Even if the clinician is interested in one gene at the time the patient presents for a consultation, the data from the rest of the genome (and other pharmacogenes) will be needed as the patients gets older, and will require other drugs in the future The challenge here is several fold: (a) interpretation of currently known variants, and ensuring the right decisions are made with respect to choosing the right drug and dose for the patient; (b) storage of that data on electronic health records for easy future retrieval and being able to act on it when the patient needs another drug; and (c) the ability to make the system dynamic so that as future drug-gene pairs are identified, the knowledge and the requirement to act on it are made available to clinicians, who will need the appropriate decision aids to enact the change. This will lead to challenges for regulators, guideline developers, companies developing decision support systems, and for the payers. Ensuring that the healthcare workforce has sufficient knowledge and skills to enable implementation of these novel approaches will also be a major challenge to all healthcare systems.

Another challenge to consider is that genomic variation will rarely be the sole determinant of how a person responds to a drug. Inevitably response to a drug is dependent on a combination of factors including:

- host factors (age, gender, weight etc.),

- concomitant medications (leading to drug-drug interactions),

- disease factors (for example renal impairment) and

- genetic factors

Furthermore, in many countries, age demographics are changing, with the proportion of people above the age of 65 years increasing. This is a good thing and highlights how advances in medicine have led to increases in life expectancy. However, this also brings with it some additional challenges for all stakeholders. People growing older typically tend to live with more than one disease (so-called multimorbidity) and are often on multiple medications (so-called polypharmacy). In the European Union, we are currently undertaking a study called ubiquitous pharmacogenomics in 7 countries [33], the aim of which is to determine how a multi-gene panel covering over 40 commonly used drugs can be utilized to reduce the burden of ADRs in a cost-effective manner in such patient groups.

Another factor to consider is that most gene-drug pair associations have been defined in

European ancestry populations. Although the same gene-drug pair associations may be important for other ethnic groups, the frequency of different variants changes with ethnicity. Thus, a variant that defines a response in one ethnic group may not be relevant in another ethnic group. This has been shown with respect to warfarin. Warfarin is still a very widely used drug; it is a narrow therapeutic index with a wide dose range – some individuals require 0.5mg/day while others may require 20mg/day, a 40-fold variation. Determinants of daily warfarin dose requirements include age, body mass index, interacting medications and several genetic factors (CYP2C9, which metabolizes warfarin, VKORC1 which is inhibited by warfarin, and CYP4F2 which is involved in vitamin K metabolism). The genetic factors account for ~40% of the dose variability, far greater than the known clinical factors. Work undertaken by us was able to show that genotype-guided dosing of warfarin was superior to standard care in a randomized controlled trial (termed EU-PACT) [34]. We subsequently went onto show that this model could be successfully implemented into anticoagulant clinics [35]. Our dosing algorithm was developed for White individuals, which represented 97% of our trial population. In the US randomized trial, called COAG [36], no difference was shown between genotype-guided care and a clinical dosing algorithm. Part of the reason for this was that only 67% of the trial participants were White, while 27% were Black and 6% were Hispanic. Of importance here is that in the Black population, the prevalence of the variants which were used to develop the algorithm (CYP2C9*2 and CYP2C9*3) is much lower than in White people, and the variants which are more prevalent (CYP2C9*5, CYP2C9*11 etc.) were not assessed.

The same issue with ancestry has also been identified with HLA genotypes and predisposition to serious immune mediated ADRs such as drug-induced liver injury (DILI) and Stevens-Johnson Syndrome (SJS) [37]. For example, the drug carbamazepine can lead to immune-mediated hypersensitivity reactions such as rash, hypersensitivity syndrome, SJS and DILI. HLA-B*15:02, which is prevalent in SE Asians population has been shown to predispose to carbamazepine-induced SJS; this is now included in the carbamazepine drug label as a recommended test. However, the population prevalence of HLA-B*57:01 is less than 0.01% in Europeans meaning that it would be of little clinical utility in this population. Instead, HLA-A*31:01 predisposes to carbamazepine-induced hypersensitivity in European populations, as well as in several other populations. Interestingly, the predisposition with HLA-A*31:01 covers a number of phenotypes (SJS, hypersensitivity syndrome, DILI) while the predisposition with HLA-B*15:02 only covers SJS. Since the beginning of this century about 30 different associations have been identified between different drugs and different HLA alleles, leading to several different types of immune mediated adverse reactions. A gene panel approach which can type for all these alleles may be one solution for the future, again a challenge for regulation and

healthcare on how best to utilize such an approach.

Finally, another challenge that we face as whole human genome sequencing becomes more and more prevalent, and cheaper, is that the majority (~97%) of pharmacogenomic variants are rare, i.e., have a minor allele frequency of less than 1%, and in fact most of them are very rare (minor allele frequency <0.1%). In pharmacogenomics, to date, we have largely focused on common variants, but it is likely that the overall variation in the activity of a particular drug metabolizing enzyme, drug transporter, receptor, and other drug targets, will be due to both common and rare variants. This will need to be considered in the future as our knowledge of the rare variants increases, including their functional consequences, and combined together with the effect of common variants.

All the issues outlined above will require the development of complex algorithms which take into account all the necessary factors allowing the clinician to prescribe the right drug. The development of other -omics technologies such as proteomics and metabolomics, and their incorporation into these algorithms represents an added layer of complexity for all prescribers, for healthcare systems, for regulators, payers and for the pharmaceutical industry. In clinical practice we already measure many protein and metabolite levels as part of routine care, but broader clinical-grade protein or metabolite panels are not yet widely implemented but are likely to be developed in the future.

To conclude, I think pharmacogenomics is an important part of precision medicine, ready for implementation in some areas as already being undertaken in many countries. However, we need to do much more work to fully delineate the causes of variability in drug response and use these factors in multi-modal algorithms to ensure patients the get the right drug, at the right dose, for the right disease, and at the right time. This will be a challenge for the whole community (clinicians, scientists, regulators, patients, etc.), but one which can be tackled by working together.

Richard Beger, PhD, National Center for Toxicological Research/FDA, USA

The workflow in metabolomics begins with describing a problem such as early biomarkers of liver toxicity or cardiotoxicity [38]. Then you develop an experimental design which may include a human study, an animal study or even an in-vitro study. The metabolomics study design will include procedures of when to collect and store samples. Whether to collect blood, urine, tissue,

fecal, media or cell samples? What are sample storage and preparation conditions?

Once the samples are collected and stored, data acquisition can be conducted. Then you move on to data processing and identification of the metabolites.

Next comes the statistical analysis -- some of these are single biomarker and some of these are multivariate biomarker type approaches. And usually, at that point, we also try to do some functional interpretation. What is the underlying biology, and what was is the mechanism behind these biomarkers? And finally, the validation of the study outcome and additional biomarker confirmation studies.

In 2017, there was a thinktank meeting and 45 scientists got together and decided to start the Metabolomics Quality Assurance Quality Control Consortium (mQACC). The goal was to engage the metabolomics community to communicate and promote the development and dissemination, harmonization and best QA/QC practices in untargeted metabolomics.

Quality assurance strategies typically include logbooks, training, temperature monitoring and SOPs. For QC there are three different parts which includes sample tracking, randomization and storage and study design. The QC samples include system suitability, blanks, internal standards and pooled samples. After the experiment is completed, you use the data collected from QC samples to identify data outliers, identify peaks, and evaluate peak acceptance criteria (i.e., m/z, retention time, peak shape). System suitability samples are used to test whether an analytical system is "fit for purpose" and working within system specifications with no contamination, prior to analysis of the samples [39]. Intra-study QC samples are typically pooled samples and they have multiple purposes. Pooled samples are used to condition the analytical system, to allow analysis of interstudy reproducibility with the same samples, to monitor, assess and potentially correct for systematic errors in measurements, during an analysis, and to optionally used to filter variables based on the linearity and occupancy [39]. Some labs will have long-term reference standards or intra-lab QC standards that are often used to evaluate special class of metabolites that they run in every experiment so they can show that over time that their data between separate studies within the laboratory are consistent. Interlab QC samples are often standard reference materials (SRM) such as those that are available from NIST and allow for direct comparisons across laboratories. Process blanks or extraction blanks are considered QC samples and are used to detect and measure contaminants that may arise from the sample processing. These signals can be removed from a study during the data processing. Blanks are used to determine carryover peaks and when carryover becomes too great the analysis needs to be

stopped.

A mQACC survey from 21 metabolomics labs found that 90 percent of the labs assessed sensitivity, mass accuracy and retention time and that 70 percent of the labs used blanks to ensure instrument stability and about two thirds of the labs reported that they assessed peak width [40]. The survey found that almost 40 percent of metabolomics labs use multiple types of QC samples and another 40 percent of those labs have a specific set of standards that they evaluate. Another 20 percent of the labs used the biological sample they are evaluation and then spike in standards.

How one prepares the samples before you inject them in the instrument can also play a big role in how you measure your metabolomics data. As an example, we studied 20 healthy volunteers, collected six tubes of blood from each and then we randomized how we processed the six aliquots. The samples processing procedures: blood at six hours at zero degrees, blood at six hours at room temperatures, plasma after 24 hours at four degrees C, and plasma at room temperature for 24 hours were compared to blood converted to plasma with no processing deviations [41]. The PCA indicates the variation in the experiments did not alter the outcome to any large extent (Figure 11). After PLS-discriminate analysis, we were able to cluster the samples based on how we processed them. The ones that showed the largest effect was the blood left at room temperature for six hours or the plasma which was left at room temperature for 24 hours.

Metabolomics in a clinical setting may augment PK, clinical chemistry, and functional imaging data. One can combine it with transcriptomics data, proteomics data or metabolomics data. The idea is to use all the data to get better decisions for the patient [42]. Metabolomics may clarify the role your genes can play a role in metabolite production. Inborn errors of metabolism may markedly effect enzyme activity [43]. Disorders in intermediary metabolism may affect small molecule production or primary energy metabolism. A single metabolomic analysis may identify 20 different diseases early on in an infant's life [44].

One of the first successes was the gut interaction in the acetaminophen metabolome [45]. In this example, we used pre-dose samples and were able to show that the p-cresol-sulfate interfered with acetaminophen metabolism to the acetaminophen sulfate in rats. Acetaminophen is converted by glucuronidation and sulfation and then it can produce the toxic metabolite NAPQI. Which gets deconjugation. In humans consuming two 500 mg/kg acetaminophen tablets, similar metabolites were observed. The p-cresol sulfate to creatinine ratio correlated with the amount of PSC acetaminophen sulfate [46]. So higher levels of p-cresol sulfate formed in the gut could lead

to lower levels of acetaminophen sulfate which also could result in higher levels of the toxic metabolite NAPQI (Figure 11).

In the case of acetaminophen, one may look for early predictive biomarkers of toxicity. When 200 mg per kilogram of acetaminophen was administered to mice, analyses revealed that palmitoyl carnitine increased at around four hours where the ALT increased at eight hours. In rats given 1,250 mg per kilogram we see the max of palmitoyl carnitine is at six hours, but the ALT maximum was at 24 hours (Figure 12). When humans are hospitalized because of an acetaminophen overdose, they often are administered NAC right away and that affects both the trajectories of the palmitoyl carnitine and ALT [47]. In a patient who did not get NAC treatment quickly, the palmitoyl carnitine increase to 24 hours and came back to control levels at about 72 hours but ALT increase at about 72 hours. So here the palmitoyl carnitine peak appears before ALT as were observed in rodents and humans.

Many challenges remain in the field of metabolomics. Metabolomics is a complex analytical process that can identify potential biomarkers of disease, nutritional status and drug toxicity. QC standards need to be tested and published. Metabolomics reporting standards are needed to be developed and universality accepted. Rigorous testing and validation are still required prior to regulatory approval for use in routine pre-clinical or clinical applications of metabolomics.

Susan Sumner, PhD, University of North Carolina, Chapel Hill, USA

This presentation is focused on using metabolomics to reveal biomarkers of opioid use disorder (OUD) and to inform nutritional intervention strategies. Evaluating metabolism and drug addiction has a long history.

Drs. Marie Nyswander and Vincent Dole conducted clinical trials in the 1960s that led to the development of the Methadone Maintenance Program. Their research developed the theory that addiction does not start with sociopathic tendency or addictive personality. However, they noticed that methadone prevented withdrawals and reduced cravings in opioid addicts, enabling them to return to normal life activities. And they noted that methadone restored normal homeostasis, and the amount needed varied for individuals. These observations led to the theory that addiction is initiated through a disruption in metabolism and results in persistent neurochemical disruption or disturbances, and that this imbalance is related to psychological disturbances reported for addicts.

The metabolic theory of addiction was explained to me by Dr. Jonathan Pollock, who is the chief of Genetics, Epigenetics, and Developmental Branch at the National Institute on Drug Abuse. The current diagnosis of opium use disorder is obtained through interview or questionnaires to determine if the patient meets certain DSM-5 qualitative criteria. These criteria include impaired control, social impairment, risky use, tolerance, and withdrawal. If one exhibits at least two of these criteria, they meet the diagnosis of opium use disorder (OUD), with the number of criteria met as an indicator of the severity of OUD.

Dr. Pollock and I discussed the need for objective biological markers that define opioid use disorder and to identify mechanisms needed for the development of interventional strategies. His interests lie in the identification of gene candidates, and my interests lie more towards assessment of nutrients that could mitigate against this addiction.

Dr. Pollock has been collaborating with the Dr. Arash Etemadi, who is at the National Cancer Institute, and Dr. Reza Malekzadeh, who is a distinguished professor of medicine at the University of Tehran. Dr. Malekzadeh is the PI of the Golestan Cohort Study that was initiated in the northeast of Iran to study factors for upper GI cancers, and one of his mentees (Dr. Reza Ghanbari) conducted research in my laboratory as a NIDA Invest Fellow. More than 50,000 of the volunteers were analyzed for opiate use and its complications, and more than 8,000 individuals reported opium use for mean duration of up to 13 years, either by ingestion or inhalation. The opium use disorder was determined using the equivalent of a DSM-5 for a subset of the subjects.

There are many clinical measures and health phenotypes in this cohort, including diabetes, cardiovascular disease, and cancer. In the subset of urine samples that we received for the metabolomics analysis; all the subjects that were included in this subset were deemed healthy with the exception that some of them had opium use disorder.

What were our study questions? The first question was - what metabolic perturbations are induced by opium exposure? For this question, we compared the metabolic profile of urine samples from 218 high opium users with the metabolic profiles of 80 nonusers (the control) [48]. The second question was - what are the metabolic markers that define an OUD positive diagnosis [49]? In this case, we compared the metabolic profiles of 138 high opium users who were diagnosed as OUD positive, with 80 high opium users who were diagnosed as OUD negative.

The goal was to determine biomarkers of OUD and to gain insight into mechanisms associated with opium exposure that could inform the development of interventions. For our study, the subject characteristics showed that the high opium users from whom these urine samples were derived had a significantly higher use of tobacco and alcohol and a lower BMI than the nonusers. Therefore, when we evaluated the metabolic differences between the opium users and the nonusers, those perturbations arise from the influence of the opium, tobacco, alcohol, and BMI. The subject characteristic data for our sample indicates that the opium users that were diagnosed as OUD positive did not have significant differences in tobacco use, alcohol use, or BMI compared with these opium users who were diagnosed as OUD negative.

For this study, we used untargeted UPLC-high resolution-Orbitrap mass spectrometry, and we also used NMR spectroscopy. For the mass spectrometry method, we used a Q-Exactive HFx system and detected ~ 5,000 features after data filtering. We compared signals for the study phenotypes of a) high opium users vs nonopium users, and b) high opium users diagnosed as OUD positive with high opium users diagnosed as OUD negative. We used univariate and multivariate statistics to determine the variable importance to projection, as well as to calculate p values and fold changes. Logistic regression modeling was also performed. The results of these analyses were recently published [48] and [49]).

Signals that were important to defining the study phenotypes were matched against our in-house physical standards library of about 2,000 compounds. Our library includes endogenous compounds from host metabolism, as well as exogenous compounds that can be derived from chemicals in the environment or foods, drugs of addiction, medications, tobacco use, or metabolites derived from ingestion of foods. We used big data analytics to annotate signals to public databases if no match could be made to the inhouse physical standards library.

Approximately 3,800 peaks differentiated (p<0.1) high opium users vs non-opium users, while ~ 712 peaks differentiated high opium users diagnosed as OUD positive from high opium users diagnosed as OUD negative (Figure 13). There were 519 peaks common to both differentiations, while 193 peaks were specific to differentiation of OUD positive diagnosis from an OUD negative diagnosis.

For the 218 high opium users versus the 80 nonuser controls, we saw signals that matched to metabolites of opium, such as codeine, morphine, and their glucuronides. These signal intensities were on the order of E-40 higher in opium users compared with nonusers (we used background signals in the controls for comparison). Based on subject characteristic data, we

expected to see elevations in tobacco-related metabolites because of higher reports of tobacco use among opium users, and signals for tobacco related analytes were elevated (E-15) over nonusers.

We also observed perturbations in the endogenous host metabolism, including neurotransmitter metabolism, Krebs cycle metabolism, one carbon metabolism, glucogenesis, lipid metabolism, and vitamin related metabolism. We also saw that phthalate signal intensities were different between opium users and nonusers. These signals could be derived from plastics or tubing that are used in opium use. Because phthalates have been implicated in obesity, diabetes, learning, and cognition, elevated phthalates could be important in OUD.

Metabolites that could be derived from parent compounds that are formed during curing or combustion of plants were also detected in the urine of the high opium users at higher levels than the nonusers. As an example, acrylamide is known to be formed during the combustion of tobacco, and metabolites of acrylamide can be detected in urine of tobacco users. Based on stratification of the data, we believe acrylamide could be formed during combustion of the tobacco in our studies, and that it could also be formed on combustion of the opium. However, we have some more research to do to confirm that it is formed during opium combustion. But nonetheless, I wanted to mention that acrylamide has been linked to cancer outcomes, and the detection of acrylamide-derived metabolites in the urine of these high opium users is important because cancer rates are higher in this cohort of opium users compared with non-opium users. The higher cancer rates could in part be due to the higher level of toxins that can be produced on curing and combustion of plant matter.

Those were some of the metabolites and pathway perturbations that were difference between the high opium users and nonusers. Now, let's look at metabolites that differentiate high opium users diagnosed as OUD positive, and high opium user's diagnosis as OUD negative. We conducted several logistic regression models. We started with a model that used the age of enrollment and route of administration as the base model (since these were significantly different in the subject characteristic between high opium users diagnosed as OUD positive or OUD negative.) The base model resulted in an area under the curve is 0.625. The Hosmer-Lemeshow goodness-of-fit test was assessed to show that this was a good model.

We then used all 712 peaks that differentiated ($p<0.1$) the high opium users who were OUD positive from high opium users who were OUD negative. This resulted in an improved AUC of 0.720. Using all 712 peaks plus the subject characteristics, the AUC was 0.95. This

demonstrates that we have the potential of using metabolomics to get a very strong biomarker signature of OUD. This model selected 16 peaks as important to the prediction of OUD, two of which matched to our in-house physical standards library related to the tryptophan and purine pathways. Six additional peaks matched to public databases, while 8 peaks did not match to our in-house library or to public databases and are referred to as unknown unknowns. Identification of unknown unknowns is a continuing effort in the metabolomics community for expansion of public databases.

We expected to see a general metabolic disruption between the opium users and nonusers. Focusing on the metabolic pathways and the perturbations, it is evident that tyrosine and tryptophan metabolism are perturbed. In addition, there are many vitamins that are cofactors for the tryptophan and tyrosine metabolism pathways. We also saw perturbations in signal intensities for vitamins (or the cofactor metabolites) between the metabolic profiles of opium users and the nonusers. These metabolic perturbations between users and nonusers could be attributed to differences in the intake or uptake of vitamins, the conversion to cofactors, or the utilization of the vitamin that could be different between the opium user and the nonuser.

These vitamins are known to drive, or serve as cofactors, in these metabolic pathways. Four of the metabolites on the tryptophan/tyrosine pathway tested significantly different between high opium users who are diagnosed as OUD positive and high opium users diagnosed as OUD negative.

We also saw general disruption in Krebs cycle metabolism including increases or decreases in amino acids that feed into Krebs cycle or in analytes within Krebs cycle. Again, vitamins that convert pyruvate to acetyl-CoA were perturbed between opium users and nonusers. This becomes very important because pyruvate is needed to produce acetyl-CoA, which is used in many metabolic pathways, including fatty acid metabolism and one carbon metabolism, which were also perturbed. Krebs cycle metabolism could very well be linked with memory and cognition because this is a cycle that produces $FADH_2$ and NADH for the electron transport chain to produce ATP. So, you could imagine that lower levels of ATP could be linked to memory and cognition. In addition, subjects that were diagnosed as OUD positive versus OUD negative, had significant differences in sugar metabolism and one carbon metabolism in fatty acids.

In conclusion, pathway perturbations related to opium exposure included an anticipated impact on tryptophan and tyrosine metabolism. We did see perturbations in vitamin-related metabolism

and that's consistent with poor nutrition or a disruption in the absorption or utilization of vitamins -observed in addicts. Vitamins are important cofactors for the major metabolic pathways disrupted by opium: including neurotransmitter metabolism, TCA cycle, one carbon metabolism, and fatty acid metabolism.

It is logical to postulate that a nutrient cocktail of vitamins or their metabolites, vitamin-like compounds (such as choline), and fatty acids may protect against metabolic disruptions that lead to addiction. This concept is consistent with the Doyle-Nyswander theory that addiction is initiated by a metabolic imbalance. We will be conducting subsequent studies   in animal models to determine if we can protect against addiction with combination cocktails of nutrients

Tobacco-related metabolites and other exogenous chemicals produced in tobacco or opium from curing or combustion or in methods that are used to intake drugs may also play a role in mechanisms of OUD by having a competition for substrates.

Metabolites were identified or annotated that are unique to an OUD positive diagnosis and could find clinical relevance on validation. Of course, the limitations of this study are that we need a larger sample size, additional opium cohorts, and additional drugs of abuse to see if that OUD pattern is the same for different types of drug abuse (Figure 13)..

David Wishart, PhD, University of Alberta, Canada

How software and databases can enable omics-based regulatory science will be described. The strengths and limitations of the three major omics technologies: genomics, proteomics, and metabolomics to assess the safety, composition and provenance of foods, cosmetics and drugs will be highlighted (Figure 14). It is noted that metabolomics offers regulatory scientists the opportunity to potentially identify and quantify more than 80,000 different chemicals in foods, 10,000 chemicals in cosmetics and nearly 3000 chemicals in drugs. Metabolomics is also able to identify more than 2000 chemical contaminants or hazardous substances known be found in these products. The breadth of coverage offered by metabolomics allows regulatory agencies to more precisely determine their chemical composition and ascertain the allowed, illicit and potentially harmful compounds that can be found in foods, cosmetics and drugs. To provide guidance on what chemicals should or should not be in these regulated products, Dr. Wishart and his team have been developing many open-access, comprehensively annotated databases containing information on all the known chemicals found in drugs, foods, and cosmetics,

including their safety profiles. These databases include DrugBank [50], FooDB *(https://foodb.ca)* and the Toxic Exposome Database or T3DB [51].

Over the last decade DrugBank has become one of the most frequently accessed references on drug compounds for pharmacists, physicians, medicinal chemists, and regulatory scientists. It contains data on more than 2,700 small molecule drugs, 1400 biologics, 130 nutraceuticals, and 6,300 experimental drugs in early phase clinical trials. DrugBank supports a wide range of text, structure, spectral and DNA sequence searching options. It also includes detailed information on known drug structures, drug names, drug ADMET (absorption, distribution, metabolism, excretion, and toxicity) data, drug target data, mechanisms of action as well as data on drug metabolites and drug metabolism. It also contains high quality, referential NMR and MS spectra of many drugs to enable their identification and quantification. To complement the rich information on drugs in DrugBank, Dr. Wishart's laboratory has developed a food constituent database called FooDB. This database, which also supports a wide range of online searching and browsing options, contains data on almost 80,000 compounds found in 730 different raw or lightly processed foods, such as fruits, vegetables, meats, oils, and common beverages. FooDB includes information not only on nutrients and micronutrients found in these common foods, but also data on thousands of food additives put into these food items. In particular, FooDB has detailed data on the names, structures, concentrations, flavor, aroma, color, health effects and referential NMR and MS spectra of these food-associated compounds. Of course, not everything in foods or drugs is there by design or intent. To capture information on the compounds that shouldn't be in food, cosmetics or drugs, Dr. Wishart's team has developed a separate database called T3DB. This fully searchable database contains richly annotated data on nearly 3000 known pesticides, herbicides, antibiotics, endocrine disruptors, and carcinogens that can be found in foods, drugs and other consumer products. Just like DrugBank and FooDB, T3DB includes information on the chemical structures of these toxic compounds along with their names, safety or toxicity data, mechanisms of action, known targets and referential NMR and MS spectra to facilitate their identification.

While our knowledge of what chemicals can be found in foods, drugs and cosmetics is quite extensive, our knowledge of the impact that these compounds have on the human body is somewhat more limited. Indeed, understanding what happens to the many chemicals found in foods, drugs, or cosmetics after they are consumed or after they are applied is often more important from a regulatory perspective than understanding exactly what is in these products. To consolidate the information known about the chemicals that should and should not be found in the human body and their impact on human health or human physiology, Dr. Wishart's team has

assembled several online databases. These include the Human Metabolome Database or HMDB [52], MarkerDB [53] and Exposome-Explorer [54]. The HMDB contains information on more than 115,000 compounds that can be found in the human body, including both endogenous and exogenous compounds. These exogenous compounds include xenobiotics such as chemical toxins, microbial products, drugs, cosmetic chemicals, and food-derived chemicals. The HMDB includes extensive information on the normal and abnormal concentrations of these compounds in various biofluids, tissues and organs. The HMDB is fully searchable and maintains detailed descriptions of all its compounds along with data on their names, structures, synonyms, metabolic reactions, health effects as well as their referential NMR and MS spectra. While the HMDB is a very general metabolomics database, MarkerDB and Exposome-Explorer are much more specialized. Both are biomarker databases that are designed to capture more detailed information about dietary, drug, cosmetic, pollutant or workplace chemicals found in the human body. Exposome-Explorer covers 918 dietary and pollution-related chemicals with more than 10,000 reported concentration values in various biofluids and tissues. MarkerDB has a similarly comprehensive dataset, with sensitivity, specificity and concentration information on 1089 chemical biomarkers including 265 exposure markers. MarkerDB also includes information on gene, protein and metabolite markers related to the physiological consequences of different dietary or chemical exposures.

While the databases described here can play a significant role in the regulation, safety and monitoring of chemicals found in foods, drugs, and cosmetics, they really only scratch the surface of what typically needs to be assessed or analyzed. Indeed, Dr. Wishart highlighted the fact that only 5% of the signals seen in mass spectrometry analyses of human biofluids or foods match to the compounds in these databases. The other 95% of the signals correspond to unknown compounds called "chemical dark matter". Many of these unknowns appear to be the product of human or environmental chemical transformations. To address the challenge of identifying unknowns, Dr. Wishart described a technique called *in silico* metabolomics (Figure 15). This is a computational method that takes what is known (i.e., the chemicals in HMDB, DrugBank or T3DB) and uses machine learning techniques to predict the reactions and reaction products when these compounds are in the body or the environment. The software tool his team has developed is called BioTransformer [55] and it is designed to rapidly and accurately predict biologically feasible metabolite structures that have gone through phase I and phase II, microbial, or enzymatic environmental processes or transformations. Tests with a number of pilot studies are showing that this *in silico* approach can help increase chemical coverage by a factor of 10 or more. BioTransformer is now being used to generate more than 5 million predicted

biotransformation products and the full collection of predicted compounds will soon be available through a database called BioTransformerDB.

Overall, I believe that of all the available omics technologies, metabolomics offers perhaps the most fruitful approach to assessing or monitoring food and drug safety. This is because metabolites lie at the interface of the genome and the environment, making them exquisitely sensitive to both types of inputs. Past limitations with metabolomic technologies, including rather modest chemical coverage and limited biological interpretation are being rapidly overcome. Indeed, through a combination of advances in software, databases, and technology (some of which were highlighted here), metabolomics could soon be routinely used in many areas of regulatory sciences.

Reza Salek, PhD, International Agency for Research on Cancer, WHO, France

To date, several initiatives have contributed to metabolomics scientific reproducibility and standardization. For example, at the European Bioinformatics Institute based in Cambridge, UK, MetaboLights was set up to capture metabolomic experimental data *(https://www.ebi.ac.uk/metabolights/)* . Similarly, in the US, NIH common funds established an equivalent repository, the Metabolomic Workbench, *(https://www.metabolomicsworkbench.org)* to capture metabolomic studies and experimental data. Another related aspect is developing standards and providing a common framework to gather, share, and reuse metabolomic information. In 2015, COSMOS launched coordinate metabolomic activity specifically to address data standardization needs [56], developing a set of standard format and reporting templates for capturing experimental metadata [57]. The standardization work followed in a subsequent project, Horizon 2020 PhenoMeNal, funded by the European Commission, (*https://phenomenal-h2020.eu/)*. PhenoMeNal aims were to create a set of tools, portals, workflows, and pipelines to bridge the way the data is stored, processed and analyzed, all in a reproducible and standardized manner. Following the PhenoMeNal e-infrastructure initiative, the FAIR metabolomics (FAIR: findable, accessible, interoperable, and reusable) principals on experimental data sharing was supported by the European Commission. FAIR pursues a broad community involvement and collaboration building across ongoing and existing efforts (*https://www.go-fair.org/implementation-networks/overview/metabolomics/)* . (Figure 16) shows an overview of the above initiative. Some of these efforts are now also coordinated through the ELIXIR program, another European initiative that coordinates efforts across existing resources to enhance compliance and harmonize FAIR data sharing and resource usage [58].

One of the activities at IARC is the development of metabolomics application in epidemiology, exposure, and exposomics, focusing on Cancer. Throughout our lives, we are exposed to many environmental and lifestyle exposures, some of which contribute to an increased risk of developing cancer. To investigate this, over the past 15-20 years, biological specimens, particularly blood plasma, have been collected and stored at IARC as a coordinator for the European Prospective Investigation into Cancer and Nutrition (EPIC) [59]. EPIC consists of multicentric cohort studies from 23 centers and 10 European countries and includes about half a million participants. EPIC was designed to investigate the relationships between lifestyle factors (e.g., diet, nutrition and environmental factors) and cancer incidence. Metabolomics is a powerful tool to investigate small molecules present in blood plasma and to investigate its role and relation to cancer development risk. Various metabolomics approaches ranging from targeted, sensitive, and quantifiable assays to broader untargeted techniques are used at IARC. Additionally, the dynamic range of metabolites is quite broad, so we need to set up several complementary analytical assays to generate metabolomics data. With the ability to exploit various chemical properties, one can expand the metabolome captured range in an experiment. Epidemiologists subsequently can analyze such datasets and investigate the role of small molecules associated with increased or decreased risks of certain cancer types. However, one of the critical bottlenecks of metabolomics remains is metabolite identification and results interpretation. Several ongoing collaborative efforts aim to facilitate this by better reporting identification [60] and building tools for interpreting the results *(http://www.metclassnet.org)*. In conclusion, various initiatives can enhance and democratize the application of metabolomics in human health and better tackle the ongoing challenge.

Hairuo Wen, PhD, National Institutes for Food and Drug Control, China

The focus will be, on the preclinical safety evaluation of chimeric antigen receptor T (CAR-T) cells. It is well known that the radiotherapy, chemotherapy, surgery, and hematopoietic stem cell transplantation are major means of cancer treatment and have effectively prolonged the lifetime of some patients. Meanwhile, the recent emergence of cell therapy brought us novel strategies in the fight with the relapsed/refractory tumors[61]. CAR-expressing T-cells recognize a variety of monoclonal antibody-specific antigens on the cell surface, and therefore, could attack human cells by activating intracellular signals of corresponding antigens. CD19 antigen is highly expressed on the surface of B lymphocytes and their progenitor cells, as well as the tumor cells derived from B lymphocytes. CD19 antigen has become the most common target for CAR-T

products for acute lymphoblastic leukemia in children and adults and has demonstrated impressive efficacy in patients with B cell malignancies. Both Yescarta and Kymriah are two anti-CD19 CAR-T cell therapies that have been approved by the U.S.FDA for the clinical second line treatment in 2017 and provided a boost of the scientific community for the development of CAR-T[62, 63]. Further, Pembrolizumab for adult and children with unresectable or metastatic tumor mutational burden-high were approved in 2020, as the first approved CAR-T treatment for solid tumors[64]. Up to the end of May 2020, there were 41 CAR-T IND applications in China, including 31 applications targeted at CD19, taking account for about three-fourths of the total CAR-T IND applications.

Meanwhile, the toxicity of CAR-T products has drawn much attention. For instance, the on-tumor toxicities, including cytokine release syndrome, tumor lysis syndromes, organ-specific toxicities, including neurotoxicity and pulmonary toxicity, and the long-term risks has been recognized as important side effects that associate with the CAR-T cells [65, 66]. At present, both the EMA and U.S. FDA have published guidelines or documents on the preclinical assessment of cell products *(https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-human-cell-based-medicinal-products_en.pdf)*; *(https://www.fda.gov/media/87564/download)*. However, there is currently lack of detailed reference for the preclinical safety evaluation of CAR-T, and the important issues in the safety evaluation of CAR-T include choice of animal model, biodistribution, tumorigenicity, and study design. For the choice of animal model, immunodeficient mice are the most commonly used model in the preclinical safety evaluation for CAR-T products for improving the prediction accuracy. The process and/or environment of CAR-T in animals should mimic its process and/or environment in the human body to the greatest extent. The xenografting mice model without host immune system, whereas, could not completely simulate the cascade reaction brought about by the CRS in the human, as well as the off-target effect. Due to the lack of immune cell, non-tumor-bearing immunodeficiency mice, which do not produce the immune rejection reactions, allow CAR-T to proliferate *in vivo* for a longer period and are more feasible in the non-target safety evaluation. The biodistribution data could provide information on the delivery, engraftment, and cell retention, distribution, viability, proliferation, persistence, and reference for dose justification and will help to interpret the observed effects. It is a commonly used method for biodistribution including *in vivo* imaging, flow cytometry detecting CAR-T positive cells, immunohistochemistry staining and qPCR. For the concerns on the tumorigenicity, it has not translated to standard carcinogenicity studies. Testing strategies may include *in vitro* testing including several endpoints: cell growth rate, cell differentiation, cells adhesion, growth factor independent growth, expression of the oncogenes, and *in vivo* evaluations of cell proliferation.

Dr Hairuo Wen provided an example on the preclinical evaluation of a CAR-T product that against the CD19, which has demonstrated as a potent antileukemia therapy for Chinese R/R ALL[67]. In this case, the severe immunodeficiency mice (NSG mice) were adopted to construct a Raji cell xenograft lymphoma model for a comprehensively evaluation on the efficacy, biodistribution and toxicity of a 4-1BB/CD3-ζ-costimulated target CD19 chimeric antigen receptor T cell (CART19) on tumor-bearing mice [68]. A total of 120 NSG mice were used for combined pharmacodynamic and toxicity status for 56 days in 96 mice of which a single dose with Raji-Luc of half million per animal, and different concentrations of CAR-T 19 including 20 million, 60 million, and 180 million per animal, respectively. The remaining mice were left untreated. Testing of Raji-Luc in mice included clinical singleton body mass, hematological analysis, humanized cytokine, and emphasized a subset counting and histopathological examinations. In addition, a single dose of 60 million CAR-T 19 was interventionally administered into 48 NSG mice, and its distribution of CAR-T 19 in different T-cells was determined using a qPCR method. As demonstrated in the results, the proliferation of lymphoma in NSG mice treated with CART19 were inhibited and the survival time was significantly prolonged. The changes in the bio-distribution and toxicity indexes related to CART19 are mostly associated with their therapeutic effects and animal model characteristics, and the effectiveness and safety of CART19 are consistent to those of the similar products as reported. In addition, the study system and techniques applied provided predictive data to support the clinical trial for CAR-T therapy. This study shed lights for the justifications on animal models, test methods and data analysis for preclinical research of CAR-T products in the treatment of hematological tumors, and a reference for the research strategies and techniques of the preclinical safety study for newly developed CAR-T product. Their research data were used for an IND application of CART19 to the National Medical Products Administration, and a clinical trial permission was granted in China in 2019.

TRACK C: Microphysiological Systems and Stem Cells as Predictive Tools
William Slikker, Jr., PhD, National Center for Toxicological Research/FDA, USA and Elke Anklam, PhD, Joint Research Centre, EU

Emerging technologies are playing a major role in the creation of new approaches to assess the safety of both foods and drugs. However, the integration of emerging technologies in the regulatory decision-making process requires rigorous assessment and consensus amongst

national and international partners in various research communities. The need for advanced approaches to allow for faster, less expensive, and more predictive methodologies is becoming increasingly clear. In addition, the strengths and weaknesses of each new approach needs to be systematically examined. In pursuit of the goal to simulate a human—at least in terms of chemical effects, safety evaluation, and the practice of regulatory science—a system of cells or tissue may be examined under strict criteria to reflect the human condition. These "human-on-a-chip" and "human organ construct" microphysiological systems (MPS) are an emerging technology that has the potential to correlate in vivo with in vitro and simulate human organ systems. Even though the use of human cells may be an enormous advantage because there is no need to extrapolate across species, there is the requirement that different cell types be characterized in terms of developmental stage and functional capacity.

These microphysiological systems have the potential to be used to 1) assess basic biology and physiology, 2) assess the pharmacology and toxicology of drugs and chemicals, 3) study organ–organ interactions, and/or, 4) be used as a human disease model. With the use of human cells there may be the requirement that different cell types interact in a three-dimensional relationship to provide predictive value for the intact human. Another important consideration for simulating human outcome is the quantification of chemical exposure. Absorption, distribution, metabolism, and elimination are features that need to be considered. For this purpose, connection of several organs-on-a-chip with well-constructed and well-tuned fluid dynamic systems, to simulate an intact human-on-a-chip, is necessary. Microfluidic control of multiple organ systems is possible, and models envisioned by bioengineers have been developed. The challenges include the requirement for each organ type to have its own specialized media while also being connected to replicate the human circulatory system. The integration of emerging technologies into the regulatory decision-making process requires rigorous assessment and consensus amongst national and international partners in various research communities.

To address these important issues, a well-respected group of experts will provide examples and advice concerning the application of emerging technologies to regulatory science. These experts include Dr. Hajime Kohima from the NIH, Japan and also Dr. Seiichi Ishida from Sojo University, Japan; Dr. Alexandre Ribeiro, from CDER/FDA; and Dr. Janny van den Eijnden-van Raaij from the Institute for Human Organ and Disease Model Technologies, the Netherlands; and Dr. Suzanne Fitzpatrick, from CFSAN/FDA, as well as Dr. Kit Parker, from Harvard University.

The second half of Track C is focused on the manufacturing and regulatory challenges related to medical devices and therapies. The field of cell therapies is growing. For example, regenerative medicine, cellular therapies, gene therapies, and stem cells can all be used to enhance the healing processes. In this respect, of course, it is very important to understand the interaction of living cells and the environment. The following experts provide important information to understand the many challenges, the progress made already, and actions to be taken: Dr. Uwe Marx from the German company, TissUse; Dr. Kyung Sung from the US-FDA; Dr. Tao Wang from the Chinese Center for Drug Evaluation; Dr. Clive Niels Svendsen from the Cedar-Sinai Medical Center, USA; and last, but not least, Dr. Ivan Rusyn from the Texas A&M University, USA.

Hajime Kojima, PhD, National Institutes of Health Sciences and Seiichi Ishida, PhD, Sojo University and National Institute of Health Sciences, Japan

The focus is on the challenge of MPS standardization on the absorption, distribution, metabolism, and excretion (AMED)-MPS Project in Japan. (Figure 17) First, the capability of MPS is demonstrated of MPS by the reproduction of the physiological situation in its culture compartment as exemplified by hepatic zonation. The results indicate that the gradient of the medium formed depending on the position in the MPS culture compartment and affected the viability of the cells. In short, the percentage of the dead cells at the outlet side of MPS is higher than that at inlet side, mimicking the observation of region-specific hepatic toxicity induced by some chemicals in in vivo study. Such phenomenon is observed in the hepatic sinusoid. (Figure 18) As blood passes through the sinusoid, hepatocytes around the sinusoid consume oxygen, nutrients, and some factors in blood, and, at the same time, they secrete waste, factors, and metabolites. These materials accumulate as blood moves to the outlet, the central vein. Changes of the hepatocyte functions according to this gradient are known. (Figure 19) Some of them are higher at the portal vein side, and others are higher at the central vein side. These differences induce the difference of hepatocyte functions between portal and central vein sides. The result of gradient formation in the MPS culture compartment indicated that reconstitution of the region-specific hepatic function would be possible by MPS. This result and those from others [69] suggest the MPS can mimic the physiological phenomena in in vitro. Based on such possibility, some pharmaceutical companies are beginning to utilize MPS for their drug development. [70] To support these activities, discussions on the MPS standardization from the regulation side have been started in the MPS project. The chart is the standard procedure of the acceptance of new test method as a test guideline in OECD. (Figure 19) There seems to be two goals in this chart, one is "industrial acceptance" and the other is "regulatory acceptance". The final goal of MPS should be to achieve the regulatory acceptance, however, the current situation of MPS is at the research

and development stage. Thus, the immediate goal of MPS is industrial acceptance. The AMED-MPS project is working to standardize MPS as a newly developed test method. National Institute of Health Sciences, National Institute of Advanced Industrial Science and Technology, and the University of Tokyo are leading this standardization activities together with Japanese leading pharmaceutical companies.

The standardization of MPS is discussed according to its configuration. In the simplified model of MPS, the medium in the reservoir bottle flows into the culture compartment through tubing. After flowing through the surface of the cell, the medium moves out and travels to the waste bottle through tubing. The pump facilitates the medium flow. The functionality of the cell used in the device is an important component of standardization. [71]. Table 1 indicates the example of the minimum requirements for live-MPS, based on the pharmaceutical user opinion discussed in AMED-MPS project. However, due to the complicated configuration of MPS, there are other points to be considered for MPS standardization, including culture compartment fabrication, tubes and bottles, and equipment assembly to keep them viable in the culture conditions. Some of the promising study has already reported from the AMED-MPS project, such as sterilization of the equipment [72] and adsorption of chemical substances to the surface of culture compartment [73]. In summary, two different criteria should be considered to enhance the MPS performance standards. One is cell criteria, and the other is the standard for material and equipment [74]. The considerations these two aspects together will enhance the MPS performance standards.

Alexandre Ribeiro, PhD., Center for Drug Evaluation and Research/FDA, USA

Microphysiological systems (MPS) are micro engineered platforms for culturing tissue- or organ-specific cells in a designed microenvironment that mimics physiological settings and hold great promise for predicting clinical effects of drugs [75]. The FDA Division of Applied Regulatory Science (DARS) aims to move new science into the FDA Center for Drug Evaluation and Research review process to close the gap between scientific innovation and product review [76]. Overall, DARS prioritizes mission-critical applied research to develop or evaluate tools, standards, and approaches to assess the safety, efficacy, quality, and performance of drugs. DARS studies the potential of using cells differentiated from human induced pluripotent stem cells (iPSCs) maintained in a physiological microenvironment for predicting clinical drug effects [77]. To date there has been limited uptake of MPS and iPSC-differentiated cells for use in regulatory decision making. To facilitate further uptake of these technologies in the regulated

areas of drug development, DARS is developing comprehensive approaches focused in ensuring the reproducibility of using cellular microsystems.

Rationale for studying hepatic and cardiac cellular systems

Research in DARS has been centered on hepatic and cardiac cellular systems since drug-related cardiac and hepatic adverse events have led to over 75 percent of safety-related drug withdrawals. The liver is also a key organ to model drug pharmacokinetics given the roles of drug transport and metabolism in this organ in regulating drug clearance and bioavailability [75, 77]. In addition, generation of toxic or efficacious drug metabolites can occur in the liver and the field could benefit from improved approaches for predicting these events from using liver systems in drug-drug interaction studies, physiologically based pharmacokinetic modelling and *in vitro* to *in vivo* extrapolation [75]. Biomarkers, or functional or toxicity mechanisms can be assayed from liver [75, 77] and cardiac systems [77, 78].

Liver systems

The liver microenvironment is three-dimensional (3D), where different hepatic cell types cells are exposed to fluid flow [75]. Liver systems have been developed to replicate these conditions for several applications and have been demonstrated to enable long-lasting hepatic properties of cultured cells. Based on the amount of publications in the field and commercially availability of systems, it is reasonable to expect the use of these systems in drug development within the coming years [75]. DARS has studied a system with scaffolding where cells were seeded, forming 3D microtissues under fluid flow [79]. Having multiple cell types that exist in the human liver can enable the evaluation of diverse mechanisms of drug effects that depend on the function of different cell types [75]. Co-culturing Kupffer cells with hepatocytes has been demonstrated to enable the investigation of the role of inflammatory factors on the effects of drugs on hepatocyte function [80]. Inflammatory settings induced by lipopolysaccharides generally consist of increased expression of cytokines in Kupffer cells and reduced functional activity in co-cultured hepatocytes. Reproducibility of results from experiments repeated in different sites with hepatocytes co-cultured with Kupffer cells in a liver MPS was demonstrated following quality control criteria [79]. The type and quality of cells to use are of interest for meeting functional demands when aiming to recreate specific properties of hepatic function [75]. Overall MPS can contain cells isolated from humans or cells from iPSCs that are reprogrammed from somatic cells. Both primary cells and iPSC-differentiated cells are studied in DARS.

Cardiac and liver-cardiac systems

Despite recent achievements in maintaining primary human cardiac tissues in culture [81], primary cardiomyocytes are not often used in MPS, being iPSC-derived cardiomyocytes the most used cell type for these systems. In contrast to primary hepatocytes, several difficulties exist in isolating primary cardiomyocytes from donors for freezing, thawing, plating, and maintaining in culture [78]. DARS research of cardiac systems has used iPSC-cardiomyocytes, which can be easily maintained in culture but have fetal-like properties that can limit their use [77]. However, the microenvironment recreated in cardiac MPS has been reported to enhance the maturity of iPSC-cardiomyocytes [77, 78] and may increase the spectrum of applications of these cells in drug development. The interconnected liver-heart system being investigated in our laboratory was developed at the University of California, Berkeley, in the laboratory of Kevin Healy [82] and uses iPSC-derived cells.

Our ongoing experiments in establishing a heart-liver connection aim to recreate clinical results where liver metabolism regulates drug effects, as it occurs with terfenadine where inhibition of its metabolism leads to its systemic accumulation to levels that affect repolarization in cardiomyocytes [83]. A diversity of cardiac systems have been developed in the recent years using iPSC-cardiomyocytes, where some consist of 3D multicellular models with recreated cardiac physiological features that induce cellular alignment and set other properties that define functional cellular maturity [78]. Cardiac MPS can also expose cells to electrical, physical, biochemical, and biological cues to match properties of cardiac tissue microenvironment. The engineered heart tissue pioneered in the Eschenhagen laboratory [84] allows assaying drug effects on contractility. Ongoing studies with this type of heart system focuses on contractile endpoints and on ensuring reproducible and stable function between tissue batches. Upon fabrication, these tissues performed as published with baseline force values around 0.2 mN and stability in beat rate and other contractile parameters (Mannhardt et al., 2016). The mechanistic effects of drugs and the EHT ability to predict clinical drug adverse effects is also being characterized.

Future directions to qualify MPS for regulatory decisions

It is expected that the use of MPS will increase in drug development in the coming years, and standardization and quality control criteria will be critical for enabling their utility in regulatory decision making, which differs from applications in pre-regulated drug development. By evaluating several systems for the same organs, DARS aims to investigate commonalities between them and differences that may benefit specific regulatory contexts of use. Multiple stakeholders in addition to FDA laboratories are necessary to translate complex cellular systems from analytical validation in a research facility or an academic laboratory into a qualification path, where different key enablers need to be followed.

Janny van den Eijnden-van-Raaij, PhD., Dutch Organ-on-Chip Consortium, The Netherlands

Among the problems facing healthcare, the absence of appropriate drugs, and drug failures in the bench-to-market pipeline have a major impact. An important reason is the lack of human model systems that recapitulate healthy or diseased organs and their function in the human body. Organ-on-Chip (OoC) technology, particularly in combination with cells derived from patients, might help to solve these problems, enhancing drug development and ultimately benefitting patients and society: better drugs, available more rapidly and personalized, cheaper healthcare, business and jobs, and fewer animal experiments. The overall result would be better quality of life at lower cost with reduction or even replacement of animal experiments.

Organ-on-Chips are 3D cultures of cells in a 'chip', a microfluidic device in which the cells interact. They are intended to represent the smallest functional structures of healthy or diseased tissues or organs outside the body. Since the dynamics of the human body can be mimicked in a controlled way, and the patient's own cells with the corresponding genetic background can be included in the chip, this technology is expected to result in important human alternatives for the current models.

hDMT, the Dutch OoC Consortium, was established five years ago as bottom-up initiative from a multidisciplinary group of Dutch scientists. The consortium consists of the hDMT foundation and scientists from 14 partner organizations, including technical universities, medical centers, and knowledge institutes, who share their expertise, facilities, and ideas as a community. Research is done by the partners and the partners are supported by the foundation. hDMT is a not-for-profit pre-competitive technology institute, with the aim to develop human OoC models and make them available to interested users for a wide range of applications through open access publication and public meetings. hDMT partners collaborate in specific projects with many companies and other private partners in the hDMT network.

An important impulse for the Dutch OoC research is NOCI, Netherlands Organ-on-Chip Initiative, an 18.8 M€ grant from the Dutch Government for 10 years OoC research. NOCI is a collaborative program of seven research groups and is coordinated by Leiden University Medical Center (LUMC). Research focuses on brain-, heart-, and gut-on-chip, on disease mechanisms and

interactions between these organs.

Collaboration is key for hDMT, in the Netherlands and beyond. In recent years, hDMT took the initiative to build an OoC network in Europe, that currently encompasses almost all European countries. It has been most encouraging that countries, including the UK, Switzerland, Scandinavia, and France, are now also linking scientists and developers who work on OoC. This is the way forward towards a European Center of Excellence for human OoC by creating strong research collaborations throughout Europe and beyond.

The community was strengthened through the H2020 project <u>ORCHID</u> (OoC in Development), that involved seven research groups in five different countries, and was coordinated by hDMT and LUMC. With the help of many OoC experts worldwide, a community was built, a roadmap developed, and awareness created on many aspects of OoC technology [85, 86] *(https://h2020-orchid.eu/final-report-orchid-available).* The community building has led in 2018 to the founding of <u>EUROoCS,</u> the European OoC Society. EUROoCS is an independent, not for profit organization established to encourage and develop OoC research, and to provide opportunities to share and advance knowledge and expertise in this field. Anyone, anywhere in the world can become a member, and can benefit from discount for the annual conference and publication in the International Society of Stem Cell Research (ISSCR) journal Stem Cell Reports as the home journal when the OoCs contain stem cells. The society is growing rapidly with many enthusiastic contributors to further development of OoC.

A central question is how OoC adoption and use can be accelerated. With the ORCHID experts, specific building blocks of the European OoC roadmap were defined. General aspects such as ethics, education and training, and dissemination and communication were also addressed. The experts stated that the dialogue between developers, regulators and end users is essential for adoption of OoC. This task was assigned to EUROoCS and immediately taken up by establishing an Industrial and Regulatory Advisory Board.

The specific building blocks of the European OoC roadmap (Figure 20) include specification, qualification, standardization, production and upscaling, and adoption, that should result in many applications in various fields. Standardization and qualification have high priority and are crucial for further development and (end user and regulatory) acceptance of the technology.

There are many ready-to-use devices from different developers, that are used to model diseased or healthy tissue for specific purposes. Sometimes a simple model will suffice, but for other

purposes, more complex models including multiple cells types, sensors, materials, and fluid flow may be required. Not all models are reproducible among users and laboratories or qualified in an independent way. Models are based on standalone devices, each with their own equipment (pumps, tubing, sensors) that cannot be connected because a standard is lacking. For this reason, the ORCHID experts encouraged moving towards standardization of OoCs at different levels to bridge the gap between researchers and end users. For devices, this could be realized via open technology platforms as a basis to build customized solutions for specific applications.

An important step in this direction is the Moore4Medical Project, funded by ECSEL, that recently kicked off with 66 partners from 12 different countries. The OoC work package, coordinated by Delft University of Technology in the Netherlands, focuses on the development of the next generation smart open technology platforms, bringing devices from different OoC manufacturers into a self-contained and autonomous multiwell plate format, fully compatible with biological and pharmaceutical workflows.

Another example is the open Translational Organ-on-Chip platform, called TOP, developed by the University of Twente in the Netherlands. TOP provides an infrastructure for automated microfluidic chip control and enables academic and commercial chip developers to transform their OoC to plug-and-play formats, analogous to LEGO blocks placed on a 'mother board'.

Besides standardization, the ORCHID experts recommended to establish independent OoC testing centers. Via these centers, models can be independently characterized with respect to technical and biological performance. In the context of EUROoCS, a European OoC Infrastructure is envisioned with testing centers for in-depth testing and qualification, ending up in independently qualified and fully characterized fit-for-purpose OoC models, with guidelines and standard operating procedures on use and applications. The data will be stored in a virtual data center and be publicly accessible. This enables end users to select models that are best suited for their applications.

Existing infrastructures in Europe can form the basis for the European OoC Infrastructure. In the Netherlands, a national infrastructure for OoC, coordinated by hDMT, is being set up with Centers of Excellence (CoE) at all hDMT partner organizations. These centers will offer different services to developers and end users from academia and industry. The iPSC & OoC Hotel at LUMC (expert in human induced pluripotent stem cells (iPSC)), and the OoC Center Twente (expert in chip technology) are now collaborating in a pilot scheme to realize the first two CoEs within hDMT, with dedicated research staff, equipment, and training facilities. This is

expected to be an important step to accelerate building of the European OoC Infrastructure and to facilitate adoption of the OoC technology.

Suzanne Fitzpatrick, PhD., Center for Food Safety and Applied Nutrition/FDA, USA

In 2017, FDA's Predictive Toxicology Roadmap was developed by senior scientists from all the FDA Centers. A Roadmap was deemed necessary because science is advancing so quickly with systems biology, stem cells, engineered tissues, mathematical modeling and this created opportunities to improve FDA's ability to predict risk while at the same time strive to replace, reduce, or refine animal testing. FDA wanted to create a roadmap for the whole agency to work together on critical activities that could incorporate these new toxicological methods into our regulatory science goals.

FDA's Predictive Toxicology Roadmap outlines six steps that talk about working together to incorporate new predictive toxicology methods in our regulatory reviews as part of our initiative on regulatory science. The first was to work together as a group. We needed to communicate across all offices. We needed to work on research that would meet the goals of our offices. We needed to make sure our regulators were trained in new technological methods before they saw them in applications. We wanted to elevate our research we do at NCTR and other FDA Centers that it would leverage these goals. We wanted to work with all our stakeholders and to try to get that feedback from our stakeholders because we saw advancing regulatory science, especially in the field of toxicology -- was the goal that we all wanted. We all wanted the same end point of approving safe and effective products, and so, we wanted to work with the stakeholder community to accomplish those goals. More importantly, we wanted to inform the Office of the Commissioner and the public as to what we were doing in this important area.

One thing we emphasized in our predictive toxicology roadmap is the critical role of the FDA regulators. We recognize that regulators should be included upfront in any new method development because we know what the questions that need to be answered. The FDA regulators can identify gaps for additional research. Additionally, we can train our regulators upfront in these new methods so the first time they see them is not in a regulatory application. So again, working together with our stakeholders is an important goal to move predictive toxicology forward for FDA.

Our plan emphasized the focus of FDA scientists, but also emphasize the important goal of

getting regulators upfront to outline regulatory questions so that tools weren't developed and then tried to fit into a regulatory paradigm. It should be the other way around.

Another related plan that FDA also contributed to is the new Tox21 Strategic and Operational Plan. Another plan that we are also very active in is the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Road Map. This one also emphasized that regulators need to be upfront. We need for regulatory science to help identify gaps and help to work together as a community towards filling in those gaps.

We have heard of the three R's: to replace, reduce, or refine animal testing. The three R's that ran through our roadmap and the ICCVAM and Tox 21 roadmaps can be called the three C's: communication, collaboration, and commitment. FDA stresses communication with all our stakeholders, collaboration with our scientific partners, and commitment for alternatives. FDA is very committed to develop new predictive alternative methods that may help reduce, refine, or replace animal testing.

We have presented these ideas to our stakeholders, and we got a lot of feedback, and one of the things that stakeholders have said is they'd like to see an implementation plan with specific goals for the roadmap and for FDA to clearly define what our goals were and what specific actions we were doing to reach those goals.   We agreed, and we have developed an agency implementation plan. We have charged a high-level cross-agency committee to carry out these goals. This committee is the FDA Alternative Methods Work Group (AMWG).

FDA stakeholders have requested that FDA make public our progress in moving towards alternative testing. They wanted FDA to tell the public what exactly we were doing to advance predictive toxicology. FDA agreed that that's important, and we have developed a public website *(Advancing Alternative Methods at FDA. 2020. https://www.fda.gov/science-research/about-science-research-fda/advancing-alternative-methods-fda)*.
to update, in real time, what FDA is doing in the alternative field.

FDA stakeholders also wanted one entry point to FDA to present their new methods so that -- we agreed that would be a good idea. In the past, FDAers would get e-mails from scientists saying, "I've got a great method. Can I present it to FDA? What can you do to help me?" So, FDA developed the FDA Webinar Series on New Alternative Methods to allow sponsors of new technologies to introduce these new technologies to FDA. More information and criteria foe acceptable methods can be located on our website. FDA also has an email address,

alternatives@fda.hhs.gov, where the public can communicate with FDA on any idea on alternatives.

One of the criteria for these accepting new methods into the webinar series is that the developer needs to identify up front what the regulatory gap or regulatory need that the method would help fill. Additionally, the developer needs to submit some data to demonstrate the approach. That's the most important thing that FDA wants to see: can your method fit into our regulatory plan. The webinar series is only internal for If any of our FDA laboratories, which also work on alternatives or other FDA components are interested in continuing the dialogue with you, we will help them set up some type of collaboration agreement, CRADA, tech transfer, or something like that to help assist in bringing all these new methods to the FDA.

The alternative methods work group or AMWG is under the Office of the Chief Scientists in the Office of the Commissioner. I'm one of the chairs and so is Dr. Donna Mendrick from NCTR. The AMWG has members from each center. We have six program centers, including, foods, drugs, biologics, devices, veterinary medicine, and tobacco. Also included is the Office of the Chief Scientist, as well as our entire field force and NCTR. They all work together to strengthen FDA's long commitment to promoting the development of new technologies and to reducing animal testing. We believe that working together as an agency is the best way to accomplish these goals, helping each other as we build competence in these new methods. The AMWG is the focal point for how interacting with our U.S. federal partners and other global stakeholders to facilitate a discussion and development of draft performance criteria for alternative methods. Our FDA website is where we're going to be putting some of our more recent publications on alternatives.

One of the first things that the AMWG is looking at are in vitro microphysiological systems (MPS). We wanted to develop agreed upon FDA terminology for MPS because references in the literature are describing this technology with different terminology and this adds to the confusion as this field advances. FDA wanted to identify upfront that if FDA addresses MPS we do it in a consistent manner. In addition, we're working on identifying partnerships to advance the MPS technology. We plan to work on, with our stakeholders, to drafting performance criteria for MPS.

MPS is a very important technology to FDA. FDA started working with DARPA on this technology in 2011 when it just seemed almost a pipe dream that we could develop these in vitro physiological systems. FDA collaborated with both the Defense Advanced Research Projects

Agency (DARPA) and NCATS to develop MPS. From the beginning of this project, FDA advised on regulatory requirements, validation, qualifications, and what reference compounds would be appropriate. In addition, FDA indicated what regulatory questions were pertinent. This was unique because it was the first time where regulatory scientists were involved at the very beginning of a very important technology, a technology with a lot of potential. We were able to identify gaps upfront, gaps, that if filled could help what we needed to regulate our products.

FDA developed a working definition for MPS and for organs on a chip. These definitions can be found on our alternative's website. FDA believes that it's important that we agree upon an FDA definition for MPS so it's clear what we are referring to when we speak on the topic. Not only are we working with our outside stakeholders, FDA is also doing internal research on different MPS systems to help us in developing performance criteria for bringing these data from these systems into the regulatory arena. FDA has developed partnerships with different companies that are developing in vitro microphysiological systems. We started with the two DARPA chips, from Emulate and from CN BIO. We are currently expanding our interactions to MPS from South Korea, and from San Francisco. This helps FDA understand this very complicated technology. We can make sure our researchers and our regulators understand this technology when we start seeing it in a regulatory submission.

To qualify any new alternative, its necessary to start with a regulatory question, which is also called a context of use. How can this new method be used to answer a question? What is the purpose of it? Then depending on the question and where it's used in the regulatory process- is it used to prioritize or to a replacement of a pivotal study- dictates on how much validation or data we need. Developers must also define the applicability domain, the limitations, and the sensitivity send specificity. FDA suggest that developers come in with one context of use. Additional contexts can be added later. Any method can be used in a regulatory applicant but if context of use and you're qualified -- and FDA agrees with you, we want to ask you for any underlying data, just like if you put in an animal study in order to support the reliability of your model.

In conclusion, I hope I've convinced you that FDA has taken steps to advance the regulatory applicability and acceptance of novel methods. We work collaboratively with all our stakeholders to develop these new tools, to answer regulatory questions, and in identifying critical priority activities and working with partnerships We hope we'll be prepared to not only meet our mission today, but also in the future.

Uwe Marx, PhD., TissUse, Germany

The focus is an overview on microphysiological systems (MPS) and their industrial adoption and regulatory acceptance. The company is a spinoff of the Technische Universität Berlin founded in 2010, pioneering a human body-on-a-chip concept and translation of human multi-organ-chip technologies into commercial use at end-user laboratories. MPS may be defined – often named organ-on-a-chip, multi-organ-chip, body-on-a-chip or human-on-a-chip - as microfluidic cell culture devices capable of emulating human biology in vitro at the smallest biologically acceptable scale and provide a historical sketch. The history of MPS started in the first decade of the century, and progressed along two main avenues: the MPS-based single tissue engineering pioneered by Linda Griffith`s labs at the MIT, USA [87] and the MPS-based multi-tissue engineering pioneered by Michael Shuler at the Cornell University, USA [88]. Subsequently, a strategy to explore the knowledge about the smallest functional units of each human organ, that time called sub-organoids or micro-organoids, to design and operate MPS for predictive drug testing was published by Uwe Marx [89]. The strategy is because the functional units are evolutionarily conserved, subject to genetically encoded self-assembly and, given the provision of a human-like micro-organoid –specific microenvironment can self-assemble in vitro too. Just to give some examples, food absorption is accomplished by millimeter-size intestinal villi. The liver lobule is metabolizing the food and are producing proteins for the body. And finally, kidney nephrons are excreting urine and re-absorbing water into the blood stream. In 2012 this concept has been detailed into a first human "body-on-a-chip" design with a downscale factor of 1:100.000 [90]. The lung-on-a-chip of Donald Ingber's labs at Wyss Institute, USA, was the first MPS to the Journal Science [91]. Together with the enrollment of a US National tissue chip program from 2011 onwards [92] that developments stimulated an exponential growth of MPS-developments in academic laboratories globally. Consequently, those academic developments cumulated in more than thousand scientific papers on MPS in 2019 with another MPS-publication reaching the Science Journal [93].

The next question is how academic MPS inventions finally can deliver benefit for patients, the highest value for our society. This requires the adoption of MPS-based assays as decision support for new drug candidates and advanced therapies by the pharmaceutical industry and their acceptance by the authorities. The value chain to the patient's bedside stretches from academia, through suppliers, CROs and the pharmaceutical industry, to the regulatory authorities. To introduce the conclusion of major players from all stakeholder groups along that value chain in a

recent transatlantic think tank for toxicology report, that no MPS-based assay data reached authorization processes and only 23 MPS-based context-of-use assays have been involved in end-user industries for internal candidate portfolio assessment as of end of 2019 [94]. These assays were based on seven different MPS-based human single tissue/organ models (blood vessels/vasculature, bone marrow, gut epithelium, lung, liver, ocular compartment, kidney epithelium) and three multi-tissue/organ-models (liver-pancreatic islets, liver-thyroid, skin-tumor). It was outlined that this is in stark contrast to the large number of academic activities which support novel discoveries and satisfy the curiosity of scientists. The main reason is that there are major obstacles at the interfaces between the stakeholders in the value chain as highlighted in Figure 21.

Firstly, the well-known replication crisis prevents a major part of academic discoveries to be reproduced in other laboratories [95]. For MPS, this is exacerbated by the fact that most systems and chips are home-grown and therefore not commercially available to other labs for data reproduction. Secondly, a qualification gap exists for tools and models between MPS suppliers and pharmaceutical industry due to the yet immature status of the supplier industry and its focus on equipment and disposable chips rather than on chip-based qualified human biological models and assays. Thirdly, within the pharmaceutical industry, a mental barrier and intellectual property issues prevent the voluntary sharing of MPS-derived assay data with the regulatory authority if they screen out the candidate already in the preclinical phase. Dr Marx emphasized that the report provides proposals which have been developed by the MPS-stakeholders at the workshop, hold in Berlin in 2019 for overcoming these three obstacles. On one hand professional suppliers developed from high-ranking academic backgrounds filling the gap of commercially available qualified MPS equipment and chips. Prime examples are TissUse, a spinoff of the Technische Universität Berlin in Germany, Emulate – a spinoff of the Wyss Institute in Boston, US, - and Mimetas with its roots in the University of Leiden, Netherlands. On the other hand, regulatory agencies are gaining experiences applying MPS-based models and assays in their labs in the context of regulatory science. The US FDA is the front runner here with the largest know how in this area and other agencies such as the Chinese National Center for Safety Evaluation of Drugs are following now as a recent paper demonstrates [96]. The stakeholder report sketched a roadmap over the next 15 years. The authors are confident that within the next five years some of the qualified MPS-assays at end user labs will produce data used in IND/IMPD application documents to initiate clinical trials. Mature MPS-based models such as TissUse's human long-term bone marrow model and advanced cell therapies will be the key drivers for that

achievement. The next level of human MPS-based models it is expected to emulate self-contained organismal pathophysiology of individual patients through interconnection of ten or more autologous organ models [97] within the next 15 years. That might trigger the use of such MPS-based personalized patient equivalents for studies to mimic Phase 1 and Phase 2 clinical trials [98].

Due to the current lack of consistent guidelines for the validation of MPS-based context-of-use assay, TissUse has implemented and annually inspected total quality management system certified according to latest DIN ISO 9001:2015 standards. Three different qualification processes that build on each other form the backbone for the establishment of valid assays within the company or in cooperation with end users. These processes are equipment qualification according to European CE mark standards, qualification of biological models based on a specific set of that equipment, and qualification of a certain context-of-use assay building on a qualified model using its specific set of qualified equipment as highlighted in Figure 21. Adherence to good cell culture praxis standards [99] is required for such model and assay qualification.

TissUse´s current equipment portfolio consists of two devices (a HUMIMIC® Starter laboratory system and a HUMIMIC® AutoLab automated higher throughput system) and three types of chips (HUMIMIC®Chip 2, HUMIMIC®Chip3, HUMIMIC®Chip4) enabling any single organ model culture and the co-culture of 2, 3 or 4 interconnected organs respectively. An on-chip micro-pump ensures surrogate blood perfusion at near to physiological tissue to fluid ratios and pulsatile flow. One has a microscopic access to the different compartments and the circulation. The equipment is commercially available for academic laboratories exploring their own biological models on the HUMIMIC® platform. For industrial end-users e.g., in the pharmaceutical or consumer products industries a qualified biological model is mandatory prior to the qualification and use of an assay fitting customers purpose. At TissUse 16 human single-organ models and 12 multi-organ models have been established so far with this equipment. On the basis of these models, (Figure 22) context-of-use assays are currently being established at different readiness levels of end-users (Figure 23).

These readiness levels consider the fact that the interests of end-users (e.g., pharmaceutical and consumer product industries, biotech and CROs) in MPS-based solutions are very different. Thus, early adaptors like to set their high qualification standards already at the model establishment stage (readiness level I) and are interested in validating the substance tests with an established assay in interlaboratory studies (readiness level II). They aim for the long-term use of

such assays for their entire portfolio in future. In contrast, a midsize biotech company might be interested in fast assessment of a new advanced therapy under development. In this case a fully established contract testing service for their particular context-of-use is required which fits with readiness level III.

A prime example of a successful human repeated-dose bone marrow toxicity assay has been established in collaboration with pharmaceutical industry partners after adoption of the underlying MPS-based academic 4-week human bone marrow culture [100] to TissUse's HUMIMIC® platform. Guided by AstraZeneca's drug safety and metabolism department a hematopoietic active human bone marrow model has been characterized and qualified for a stable and reproducible performance over 8 weeks. Process costs have been decreased and a repeated dose scheduling assay has been established to assess lineage-specific toxic effects. Technology transfer and inter-laboratory studies completed the qualification process. Success of the qualification and attractiveness of the model stimulated other pharmaceutical companies to engage in portfolio candidate testing with the same underlying biological model. Once established and accepted at that highest level of qualification, the added value of such human organ models feeds back into high-ranking academic labs. The bone marrow model, for example, attracted the interest of George Duda's labs at the Julius Wolff Institute of the Charité, Germany. They explored the bone part of the HUMIMIC®-based bone marrow model by modifying matrix components and the stimulation scheme for exposure to hip implant-associated dissolved Co and Cr at clinically relevant concentrations. This led to direct cytotoxic effects and verified binding of Cr to inter-trabecular bone matrix found in patients [101]. Such feedback loop making industrially qualified models and assays available to academic research and discovery, in turn, solves the problem of reproducibility at this level.

Kyung Sung, PhD., Center for Biologics Evaluation and Research/FDA, USA

The potential applications of microphysiological systems in characterizing regenerative medicine cellular products is the focus. As a regulatory unit within FDA, Office of Tissues and Advanced Therapies (OTAT) regulates many different types of products, including gene therapy products, stem cell, and stem cell-derived products, product for xenotransplantation, functionally mature

and differentiated cells, therapeutic vaccines, blood and plasma-derived products, and combination products that include engineered tissues and organs, medical devices, and tissues. It was mentioned that the field of cell therapy is rapidly growing, as can be seen from the number of Investigational New Drug applications (INDs) that OTAT/Division of Cell and Gene Therapies (DCGT) received in recent years (*https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products)*. It is interesting to note that there are a few licensed cellular products derived from functionally mature and differentiated cells and other than cord-blood based products, that there are currently no approved products based on multipotent and pluripotent cells in the U.S. market. This is partly due to the complexity and heterogeneity of the cellular products, which pose several manufacturing and regulatory challenges. In addition, cells are typically exposed to various conditions and exogenous factors during manufacturing processes, making the cells functionally different from their original states. This emphasizes the need for new methods and quality attributes to reliably predict the biological functions of manufactured cellular products.

Two proof-of-concept studies that used microphysiological systems to quantitatively evaluate the regenerative capacity of multipotent stromal cells (MSCs) are shown. MSCs are an attractive cell source for cell therapy because the cells can be harvested from various tissues sources such as the bone marrow and fat tissue, and they can be differentiated into bone, cartilage, and fat. In the first example, morphology analyses of 3D MSC aggregates, which were cultured in chondrogenic condition for 21 days, were performed to evaluate chondrogenic capacity of MSCs [102]. The 3D aggregate platform was chosen because it recapitulates critical cell-cell contact that is required for proper condensation and determination of MSCs [103]. MSC preparations from eight different donor sources at two different passages (early P2/P3 and late P5) were tested to evaluate donor and passage dependency. Figure 24 shows representative images of MSC aggregates in the cultures, showing different growth patterns of MSC aggregates from early and late passages. Twenty-one days of morphology analyses showed that all the MSCs tested had an initial size decrease from day 1 to day 4. Furthermore, most of the early passage MSCs showed recovery of MSC aggregate size after day 4, whereas most of the late passage MSCs did not show similar size recovery. When the team evaluated the histology, the ones that showed more size recovery showed higher disposition of the cartilage-associated extracellular matrix such as sulfate-GAG and collagen (Figure 24). They then examined the correlation of the size features with synthetic activity in chondrogenic gene expression. The data suggest that the functional matrix accumulation, but not the chondrogenic gene expression, correlated strongly with aggregate morphology. Overall, the study provided a method for the early estimation of chondrogenic differentiation capacity, which may have important implications for the

manufacture of high-quality MSC products.

In the second example, a 3D microfluidic co-culture platform that was used to investigate vasculogenic potential of MSCs is described. In addition to the direct differentiation of MSCs into cartilage-forming cells, MSCs are known to produce several paracrine factors that could stimulate angiogenesis or vasculogenesis. The proposed clinical applications for these paracrine interactions include wound repair, immunomodulation, and ischemic reperfusion. A compartmentalized co-culture platform [104] was used to measure the MSC paracrine effect on the stimulation of the vasculogenic network formation and the influence of manufacturing parameters, such as donor and cell passage. In the system, endothelial cells (HUVECs) are loaded in the center channel, and stromal cells were encapsulated in fibrin hydrogel injected in the side channels. After the qualification of the microfluidic device, a small screen was conducted using MSCs from four different donor sources at two different passages: passage three and passage five. Data suggests that the vasculogenic potential of MSC trophic factors appears to be influenced by manufacturing parameters, such as the donor source and passage number (Figure 24).

To conclude, methods that could have predictive value for biological activity and thereby lead to improved cellular product characterization are described. In addition, different microphysiological systems could be used to understand the influence of various manufacturing parameters on the quality of cellular products and to identify the quality attributes that may have an impact on the safety and effectiveness of cellular products.

Tao Wang, PhD., National Medical Products Administration (NMPA), China

An overview from the regulatory perspective of Anti-SARS-Cov2 drug and vaccine development in a public health emergency is provided. Since the outbreak of the COVID-19 public health emergency in early 2020, Center for Drug Evaluation (CDE) has been committed to the development of anti-SARS-Cov2 drugs and conducting review in accordance with the *Vaccine Administration Law* and the *Special Review and Approval Procedure for Drug Registration* issued by former State Food and Drug Administration, order No 21. CDE's work mainly involves the establishment of regulatory standard system and encouraging regulatory measures.

1.Establishment of Regulatory Standard System: CDE has organized experts to discuss and establish a regulatory standard system for IND application and NDA application, including

specific requirements on pharmaceutical, non-clinical and clinical data.

CDE has published five relevant guidances on its website *(CDE Guidances. 2020. http://www.cde.org.cn/news.do?method=largeInfo&id=137cd502f7584a3a)*. Chinese.

Guidance for Development of Prophylactic Vaccines for Novel Coronavirus (Interim), Guidance for Pharmaceutical Studies of Prophylactic mRNA Vaccines for Novel Coronavirus (Interim), Guidance for Non-clinical Efficacy Studies and Evaluation of Prophylactic mRNA Vaccines for Novel Coronavirus (Interim), Guidance for Clinical Studies of Prophylactic Vaccines for Novel Coronavirus (Interim), Guidance for Clinical Evaluation of Prophylactic Vaccines for Novel Coronavirus (Interim). CDE has also drafted 2 guidances on antiviral drugs which are now under discussion, namely, *Guidance for Clinical Trials of Antiviral Drugs for Prevention of COVID-19* and *Guidance for Clinical Trials of Antiviral Drugs for Treatment of SARS-Cov-2 Pneumonia*.

The guidances mainly cover non-clinical and clinical aspects. Regarding non-clinical aspect, for small molecular drug, it's necessary to define the mode of action and predict the in vitro antiviral activity (CC50, EC50, SI), to predict in vivo antiviral concentration and obtain data on virus load, lung histopathology, symptom improvement, mortality in animal infection models (such as ACE2 transgenic mice, non-human primates) before entering clinical trial; for monoclonal antibodies, the sponsor is required to submit data on characterization, antibody binding sites, in vitro antiviral activity (neutralization activity, EC50), before entering phase II clinical trials, animal infection model data should be available; for prophylactic vaccines, the sponsor is required to submit data on pharmacodynamic research such as vaccine immunogenicity and in vivo protection.

Regarding clinical aspect, for the purpose of promoting safe and effective drugs to market as soon as possible, the requirements are flexible with support to adaptive design under the premise of ensuring the safety of subjects. For new drugs for the treatment and prevention of COVID-19, CDE has established standards for the overall design, efficacy endpoints, and research cycle of confirmatory clinical trials for different target populations; for prophylactic vaccines, the protection effect of the target population is required to reach over 70%, or at least 50% (point estimate), the lower limit of 95% CI should be no less than 30%, and it is best to provide protection for one year or more, but at least 6 months. Establishment of Encouraging Regulatory

Measures: To encourage the development of anti-SARS-Cov2 drugs and Covid-19 vaccines, before submitting IND application, the applicant is encouraged to fully communicate with CDE. We allow rolling submission and conduct review simultaneously. If the above fundamental requirements are met, the applicant shall submit IND application, and CDE will conduct the review through the special approval procedure and authorize approval for conducting clinical trial. During clinical trial, the applicant shall continue to carry out pharmacovigilance and report to CDE on a regular basis. When key confirmatory clinical trial results are obtained, or the mid-term analysis demonstrates evidence of expected positive effect, the applicant shall submit application for NDA and special approval, and CDE will decide whether to approve production based on the established standards after a comprehensive evaluation. During the entire development process, whenever encounters problems, the applicant is encouraged to communicate with CDE. After the drug is marketed, the holder of the drug marketing authorization shall be the main body of responsibility to continue pharmacovigilance study and fulfill post-market requirements (Figure 25).

Outlook: At present, CDE has approved several drugs to carry out clinical trials, including small molecules, monoclonal antibodies, and prophylactic vaccines, some of which have entered critical phase III clinical trials. Under the situation that the epidemic in China is well controlled, we are facing a huge challenge of how to conduct a phase III clinical trial. The next step is to strengthen international cooperation in global clinical trials and harmonize the regulatory requirements of national regulatory agencies. At the same time, we need to strengthen cooperation with institutions, the public, industry, and academia.

Clive N. Svendsen, PhD., Cedars-Sinai Medical Center, USA

Induced pluripotent stem cells (iPSCs) can generate neurons in the culture dish from patients with specific neurological diseases, and as such represent a powerful new human model for these devastating disorders. Here, is described how one can combine iPSCs and microfluidic organ chips to enhance regulatory decisions for two neurological diseases - amyotrophic lateral sclerosis (ALS) and Parkinson's Disease (PD).

ALS involves the death of both upper motor neurons in the motor cortex that project to the spinal cord and lower motor neurons in the spinal cord that project to the muscles. Muscle weakness presents in early stages of disease and ultimately paralysis and death occur typically within 4-6

years of diagnosis. For sporadic ALS, which encompasses approximately 90% of cases, there is no known cause and no successful treatment. The remaining about 10 percent of cases have genetic mutations, which are being targeted by various drug companies.

In PD, dopamine neurons in the substantia nigra die and loss their projections to the striatum. The reduced dopamine leads to debilitating changes in movement. There are also early effects on gut function and other peripheral systems. As with ALS, the majority of cases are sporadic with no known genetic cause, though there may be an environmental component in some cases. Only about 10 % of cases are familial with known mutations in genes like *LRKK2* and *SYNUCLEIN*. There are therapeutics that can affect levels of dopamine and other brain chemicals to alleviate some symptoms, but there is no treatment that halts disease progression.

Animal models for ALS and PD have limitations and there are no human-based models to test toxicity or efficacy. The need for new models for neurologic disorders could be addressed by patient-derived iPSCs. Somatic adult cells, for instance from the blood, can be reprogramed using a cocktail of transcription factors, which takes them back to something like an embryonic stem cell [105]. iPSCs were initially considered for transplantation as an autologous cell product. However, this is not yet mainstream, with only a few clinical trials in the world. The bigger use is for personalized medicine whereby an iPSC line can be created from an individual patient. The iPSCs can be differentiated into the different disease-relevant tissues in order to learn about disease mechanisms and develop drug therapies, which my lab recently did using iPSCs derived from young onset PD patients [106]. However, one limitation is that iPSCs are at an immature developmental stage, which may hinder the study of late-onset diseases. Additionally, the cells are often maintained in two-dimensional (2D) monoculture, which does not recapitulate human physiology. As such, about 5 years ago, my laboratory began to use a microphysiological organ chip system to permit 3-dimensional (3D) culture of multiple cell subtypes to mimic conditions of human physiology *in vitro*. Cells are seeded onto the organ chip (from Emulate Inc), which has two channels separated by a porous membrane.

We wanted to combine this powerful organ chip technology with iPSCs for regulatory science, in particular related to neurological diseases. This required several steps, that included creating a blood brain barrier (BBB) and then culturing cell subtypes relevant to ALS and PD on a chip. The BBB is a critical component of the central nervous system that allows the flow of nutrients from the blood into the brain and that protects brain cells from potentially harmful substances found in peripheral circulation. Advances in iPSC and organ-chip technologies now allow us to improve our knowledge of the human BBB in both health and disease [107].

We created a BBB chip with iPSC-derived brain microvascular endothelial cells (BMECs) and neural cells [8] (Fig. 26). BMECs are the vascular cells that line the blood vessels of the brain, and permit transport of drugs into the brain. These cells are typically difficult to create *in vitro*, but can be efficiently generated using the iPSC technology. We call them BMEC-like because the iPSC-derived cells don't entirely recapitulate brain endothelial cells, but they have the correct characteristics mechanistically to study the BBB. Using the same iPSC lines, neural progenitors, neurons, and astrocytes were generated and seeded to create a brain channel. The BMECs are able to block molecules, like Dextran, from permeating into the neural side.

The iPSC and organ chip technologies can enable disease modeling and personalized medicine. For instance, we modeled the childhood syndrome Allan-Herndon-Dudley syndrome (AHDS), in which a genetic mutation leads to moderate to severe intellectual disability and problems with movement. By differentiating iPSCS into BMECs and neural cells and culturing them on an organ chip, we confirmed a lack of thyroid transport across the BBB in AHDS [105]. Perhaps, most exciting in the case of regulatory sciences is that blood can be flowed through the BMEC-lined channel in order to test transport of compounds through the BBB from the blood side to the neural side. Further, a BBB derived from iPSCs of Huntington's disease patients showed an increased penetrance of molecules suggesting a deficit in the BBB in disease [109]. This BBB data correlates with the *in vivo* patient data, demonstrating the strength of iPSC and organ chip technologies for disease modeling.

This approach has advantages for regulatory science as a drug's ability to cross a tissue barrier can be tested from a patient's blood rather than using culture media. You can also assess toxicity on the neural cells after a drug flows from the blood through the BBB. Collectively, the results confirm that this BBB chip can recapitulate how proteins and drugs can get from the blood into the brain and in health and disease, how this can affect brain cells. This information could be a basis for new regulatory decisions.

Motor neurons are now being produced from iPSCs derived from patients with ALS[110]. These have been seeded onto chips along with BMECs from the same iPSC line (Fig 27). In doing so, a 3-D reconstruction of the spinal cord where the motor neurons in ALS undergo the very earliest stages of the disease are formed. One can visualize live neurons firing in the ALS chip and can study how this is affected when a drug is applied in the endothelial channel. The endothelial cells actually enhanced the maturation of the brain cells, emphasizing the importance in biology and in regulatory science to have mixed physiological systems and not just monocultures that do not

provide the whole story. Similar chips can be devised using iPSCs from patients with PD. Dopamine neurons, microglia, and BMECs for the BBB, can be differentiated so that one has essentially modeled Parkinson's disease on a chip. These new models of disease can be used for analysis of drug development to test drug toxicology and efficacy as well as drug delivery through the BBB (Figure 26).

The importance of controls for regulatory science is of great importance. If you use cells from a middle-aged individual as a healthy control, it is possible that they could develop a variety of diseases in a few years. To circumvent this, we created a set of control iPSC lines from a large cohort of people in Scotland who survived to at least 82 years old with no co-morbidities or brain disease. To date, 24 lines have been generated from this Lothian cohort, with a paper recently published on how the lines were made and details of each subject[111].

Another important aspect is stability. It is better to generate iPSCs from cells that have not been expanded *in vitro*. For instance, peripheral blood mononuclear cells (PBMCs) taken directly from the blood. iPSC lines made from cells not expanded *in vitro* remain highly stable. They only show abnormalities at the low rate of about five percent, compared to expanded fibroblasts that have a transformation rate of around 24 percent in culture. The same is true for embryonic stem cells that frequently develop karyotypic abnormalities. Stability of cell lines is a very important point for regulatory science, because if they vary over time or pick up abnormalities, that may affect the ability to use them to reliably predict drug effects and toxicity. The PBMC-derived control lines are available through the Cedars-Sinai iPSC Core (Figure 27).

For the last six months progress is being made on COVID-related issues, and of course a lung chip would apply beautifully to the study of COVID. My group has now created a lung chip and, in collaboration with UCLA, we are adding the SARS-CoV-2 virus to the lung chips to assess the effects and ultimately, we can test drugs to prevent infection. One idea is to use antisense oligonucleotides to knockdown the virus itself or critical receptors on host cells in order to get reduced infectivity. We are also investigating the role of SARS-CoV-2 virus on heart cells, and recently showed that iPSC-derived cardiomyocytes undergo apoptosis and stop beating at 72 hours after infection[112]. This provides a model to elucidate infection mechanisms and potentially cardiac-specific antiviral drug screening platform.

We recently published a review of the field, with the idea that 2D cultures provide high throughput and may be good for drug screening, but the physiological relevance may be low (Fig. 28) [113]. As one uses more engineered and valuable iPSC and organ chip systems, the

physiological relevance increases, but the throughput is reduced. We suggest, therefore, a tiered approach to regulatory science – starting with 2D-based screening, but then proceeding with more biologically-relevant 3D human models (Figure 28).

To conclude, the combining these technologies will provide the next generation of models for regulatory authorities to use. The questions that we have remaining include how complex do these models have to be to provide reliable toxicology and drug effect predictions in humans? It may be that while 2D cultures can provide an outcome, a more complex and physiological model may be required to definitively answer the question. It ultimately comes down to validation. We can make the best model in the dish or in a chip, but we need to validate whether what the model the predicts, be it toxicology or drug effect, is a good predictor of what happens in human disease. Then, and only then, can we really incorporate these exciting technologies into mainstream use in regulatory science.

Ivan Rusyn, PhD., Texas A&M University, USA

The topic of testing reproducibility of microphysiological systems (MPS), or tissue chips is explored. The path to wide acceptance of new technologies, such as MPS, in biomedical and regulatory sciences lies through the studies that establish the reliability, robustness and reproducibility. The need for systematic evaluation of the MPS before they can be used for environmental health and drug safety decisions was addressed by several prior meetings, most notably by the meeting held by the National Academies of Sciences that occurred just over six years ago, in 2014 *(https://www.nationalacademies.org/event/07-21-2014/the-potential-of-the-tissue-chip-for-environmental-health-studies-workshop)*. Over the past decade, there has been a tremendous amount of progress made in developing MPS devices modeling various organs and connecting some of the organs together [114]. However, MPS introduce numerous technical challenges as they require specialized training and new equipment, because they are much more complicated than traditional cell culture models. In parallel with new advances in biomedical engineering, the next frontier is to not only make these devices accessible to the wider toxicology

community but also to bring together the developers and end-users to better understand where these devices can be applied for making decisions about human health [115].

One path to bridging the gaps in MPS application was to establish consortia that would share knowledge, develop best practices, and align protocols on how these devices can be translated from the developers to users and then, eventually, to the decision-makers[116, 117]. One such consortium was established in 2016 at Texas A&M University with funding from the National Center for Advancing Translational Sciences (NCATS), this collaborative established partnerships with over 20 academic centers around US to bring their MPS technology to the Texas A&M University Tissue Chip Testing Center (TEX-VAL) for examining of reproducibility and establishing the contexts of use for each device based on the feedback from the potential end-users in the pharmaceutical industry and government agencies.

Several case examples of the experiences that TEX-VAL Center had with tissue chip testing are explored. The process was standardized into a workflow (Figure 29) that was followed for each of the 20+ platforms tested in 4 years since 2016. Execution of the material transfer agreements, establishing the right procedures in the testing laboratory, and ensuring that proper equipment and analytical methods to work with very diverse MPS are all required steps before replication of the key experiments that can demonstrate replication of the published studies can be achieved. The ultimate goal of these experiments is to better understand how these MPS can be used and to demonstrate to prospective users the value added for both toxicology research and for drug development.

The first example that was presented included the proximal tubule on the chip MPS [118], a commercially made device about the size of a credit card. The device has two channels and the chamber inside of it. For the proximal tubule chip, a tubule is made through the middle of the chamber by filling it with extracellular matrix and letting it polymerize. Renal proximal tubule epithelial cells (RPTEC) are injected into the tubule and they attach to the sides of the channel when the media is perfused through the device. At TEX-VAL, experiments were first conducted over seven days to understand what the binding of different chemicals and drugs to the device is. Second, previously published experiments [119] were replicated in experiments over 24 days. Finally, we have expanded the context of use for this device by testing various chemicals and drugs [120]. In parallel to the MPS, traditional two-dimensional cultures were also tested. In addition, both human primary RPTEC isolated from human kidneys were compared to immortalized RPTEC that are widely available albeit not as physiologically relevant as the

primary cells. The first question that was addressed was model reproducibility. While there were no significant differences in cell viability and morphology between culture conditions in 2D and 3D or between cell sources, it was found that the primary cells have much lower secretion of KIM-1, and the commercial cells have a much higher secretion of KIM-1. The gene expression profiles of different cells under these different conditions were examined and it was found that close similarity between commercial cells and primary cells exist even though their gene expression profiles were not identical. Important differences between cell sources and culture conditions were observed when known nephrotoxic compounds were tested, showing that MPS offers a more physiological response. These studies showed that the cell sourcing is critical for robustness and replicability of this MPS. However, it was also found that for some of the phenotypes, the more complex MPS may not be needed as two-dimensional cultures were quite relevant and giving us the results that were expected. This is important because the throughput of 2D and 3D models differs considerably (multiple 384/96-well plates vs 24 MPS systems per incubator).

The second example was the liver MPS developed at the University of Pittsburgh [121]. In this model, human hepatocytes are combined with various non-parenchymal cells. This model was shown to be physiological insofar that the cells were producing urea and albumin at much higher rates than two-dimensional cultures and for up to 14 days. In the original study, the developers have shown classical responses to traditional hepatotoxicants. While this model is quite technically challenging as it requires multiple cell types, it was successfully transferred to TEX-VAL and tested for both inter-laboratory reproducibility and for comparisons of human primary hepatocytes to induced pluripotent stem cell (iPSC)-derived hepatocytes [122]. In the reproducibility experiments, both the University of Pittsburgh and TEX-VAL obtained the same batch of primary human hepatocytes from a vendor. It was found that in both laboratories the average albumin and urea production over 10 days in culture in the MPS were very similar between the two laboratories and close to human liver levels. Another important reproducibility experiment involved testing the ability of these MPS to metabolize drugs. In both laboratories, terfenadine was metabolized, and fexofenadine was generated; while the rates were somewhat different (by a factor of 2), overall a very similar response was observed. Finally, the MPS was tested with primary and iPSC-derived hepatocytes treated with several hepatotoxic compounds. Also, both two-dimensional and three-dimensional cultures were tested. These experiments showed good reproducibility of the model between laboratories and demonstrated that long tern (over 14-days) culture of both primary and iPSC-derived hepatocytes is achievable and their

function far exceeds that of 2D cultures. Drug-induced responses were also similar between two cell sources and were more like clinical effects when MPS was used.

In conclusion, MPS will not be used in isolation, just like any other device or method in biology or in toxicology. And ultimately, the developers and end-users of MPS need to carefully consider the potential for regulatory decision-making with the MPS-derived data. The regulators remain to be convinced that these devices are not only relevant for human health decision-making but are also reproducible and accessible to a wide range of prospective users.

Track D: Bioimaging, Serguei Liachenko, PhD., National Center for Toxicological Research/FDA, USA and John C Waterton. PhD., University of Manchester and Bioxydyn Ltd, United Kingdom

In vivo imaging is not a new concept. Already by the beginning of the last century radiography was becoming an important tool in medical research and patient care, and by the 1950s radiographic biomarkers were used to assess disease progression and response to therapy, particularly in oncology and arthritis. Imaging techniques now provide many powerful tools in drug development. In preclinical toxicology they can aid investigating distribution of labeled compounds throughout the body, determining pre-existing pathologies in subjects selected for toxicity or efficacy studies, or monitoring pathologies throughout the in-life phase of such studies.

In clinical drug development they can provide biomarkers for toxicodynamic assessment, monitoring, or to predict harm from therapies in specific patients. In the regulatory sphere, imaging biomarkers are often used in drug labeling to guide safe prescribing. The BEST resource *(https://www.ncbi.nlm.nih.gov/books/NBK326791/)* defines a biomarker as a *"defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or biological responses to an exposure or intervention, including therapeutic interventions"*. Biomarkers can be measured with various technologies, including radiographic (i.e., imaging) technologies. BEST defines several categories of biomarker, including safety.

Why would an investigator employ imaging biomarkers for drug safety assessment? Imaging is informative for most organs of interest to drug safety scientists and toxicologists, notably the liver, kidney, brain, lung, and heart. Unlike circulating biomarkers, which provide a global whole-body average assessment, imaging inherently assesses focal damage, including early and

minor pre-toxic changes. Importantly in the management of drug safety, imaging allows follow-up, to assess whether toxic changes are progressive or reversible.

The use of imaging biomarkers in drug safety is quite well-established, and they are used routinely to avoid adverse drug reactions in trials, as well as for patient management in healthcare. Many imaging biomarkers have been evaluated as part of the FDA biomarker qualification program *(https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/biomarker-qualification-program)*, and some have been used by FDA and other agencies as surrogate endpoints. In a few cases, imaging biomarkers been recommended or adopted as companion diagnostics. Table 2 lists some representative imaging biomarkers for drug safety, many of which will be familiar to the drug developer.

The five following authors provide an overview of the potential and actual use of imaging techniques, including magnetic resonance imaging (MRI), positron emission tomography (PET) and computed tomography (CT) in drug R&D and regulatory safety science across multiple therapeutic areas. These leading experts from the pharmaceutical industry, contract research organizations, academia, and government agencies described examples on how such technologies can support regulatory decision making and have assisted in balancing treatment benefit with risk of harm.

Serguei Liachenko (NCTR/FDA, USA) describes the unmet need for preclinical assessment of neurotoxicity liabilities, and the development and validation of MRI biomarkers of neurotoxicity in the rat. Ira Krefting (CDER/FDA, USA) discusses FDA's approach to assessing gadolinium-based contrast agent safety in the light of safety concerns that have arisen with these diagnostic drugs post-approval, notably nephrogenic systemic fibrosis, a serious and sometimes fatal adverse effect, and more recently the observation of gadolinium retention in brain long after dosing. John Waterton (University of Manchester and Bioxydyn Ltd, UK) discusses the development and validation of translational imaging biomarkers of harm or lack-of-harm, and the use of public-private partnerships for imaging biomarker validation in drug safety assessment. Timothy McCarty (Pfizer, USA) discusses the role of imaging to assess drug safety and described several real-world examples of decision-making in drug development.
Yan Liu (Median Technologies, France) discusses the development of AI-based imaging biomarkers for precision medicine, and the challenges in developing and introducing this new class of imaging tools.

Serguei Liachenko, PhD., National Center for Toxicological Research/FDA, USA

An update on the development of nonclinical imaging biomarkers of neurotoxicity is presented. The current practices for testing of neurotoxicity in a lab are more than a century old and are mostly based on microscopical examination of the brain tissue. This involves brain dissection, fixation, slicing, staining and assessment by highly trained and certified specialist. Current best practice calls for evaluation of a limited number of selected areas, generally in the range of 7 slices per brain. However, even in such a small animal as a rat, the brain is complex and diversified such that a limited number of sampled tissue slices may drastically increase the chance of missing lesions depending on their size and distribution [158]. Preclinical failures to identify neurotoxicity using this approach have led to higher frequency of neurotoxicity incidents discovered at clinical stage [159], which presumably led to a higher cost and a lower patient safety in drug development.

To improve the prediction of clinical neurotoxicity we need to modify our current testing paradigm and one of the promising approaches is non-invasive magnetic resonance imaging (MRI). It offers a variety of methods including anatomical, diffusion, perfusion, relaxometry imaging, as well as spectroscopy, which could be used to assess morphological and functional tissue characteristics including cell integrity and density, edema, vacuolation, blood flow, microstructure, and neurochemistry[160].

Quantitative proton T_2 relaxation mapping was chosen for development of the biomarker of neurotoxicity. Specifically, T_2 relaxation is the decay of transverse magnetization that was acquired by protons during exposure to resonant radiofrequency pulse in the presence of magnetic field. As most of the protons in the biological tissues belong to water, the T_2 relaxation changes will reflect the perturbations in the tissue water quantity and quality. Thus, edema, change in cell number and integrity, temperature, energetics, blood flow, and other factors, related to neurotoxicity may result in changes of T_2 relaxation. The measurement of T_2 relaxation is easily implemented using "off-the-shelf" methods within reasonable time limits. Figure 30 shows the example of the T_2 map of the normal rat brain, where the corresponding T_2 values derived from signal decay curves using exponential fit are easily distinguishable between different tissue types [161]. These maps are quantitative as each voxel represents apparent T_2 relaxation time of the protons expressed in time units (milliseconds). This quantity depends only on the tissue water properties and the strength of the magnetic field and thus its measurements should be reproducible between different imaging labs/sites.

The biological variability of T_2 relaxation of the normal rat brain at 7 Tesla is low, which allows to detect 5% changes with statistical power of 0.8 in most anatomical regions using only 3-5 animals per group. However, in a proof-of-concept experiments using neurotoxicants at a sub-LD50 doses the changes in T_2 relaxation times were up to 3-fold from the control values (Figure 31) [162].

Each tested neurotoxicant created its own pattern of damage, depending on the mode of action. The damage of both gray and white matter could be seen in the same map, while with histology it would require using different stains, specific to tissue type. Also, the same subject could be imaged several times over the course of the lesion development, which could establish the time course and the potential reversibility of neurotoxicity (Figure 32) [163].

Quantification of T_2 MRI maps requires good co-registration of the images to the common anatomical space. Then T_2 biomarker could be calculated by statistical mapping as the volume of tissue with positive changes from baseline or controls or as an average value in each anatomical region after segmentation [143, 144, 161].

For biomarker to become a useful drug development tool it must be accepted either by widespread use in scientific community or by a special qualification process developed by regulatory entities including FDA and EMA [164]. One important step in such qualification is establishment of the biomarker performance against current "golden standard" method. Such performance could be described as biomarker sensitivity (the rate of correct identification of toxicity) and specificity (the rate of correct identification of absence of toxicity). To establish the global sensitivity and specificity of T_2 MRI biomarker, ten well-known neurotoxic compounds were used at doses ranging from ~ 3% to 80% of the LD50[162]. Saline-treated group was included as a control. T_2 MRI was performed at the end of observation (from 2 to 21 days, depending on the compound) immediately before brain fixation. Neuropathology was performed using amino-cupric silver staining in vast number of brain slices (80 slices per brain). All subjects were classified into true responders and non-responders based on the report provided by certified neuropathologist. Responders were all subjects with any type of pathology in any brain area. Then the same subjects were classified separately into MRI-responders and MRI-non-responders, based on T_2 MRI signal. MRI responders were those subjects in which averaged T_2 value was higher than in control group or baselines in any brain area (except for cerebrospinal fluid, which was excluded from analysis). Then the numbers of true positives (TP, true responder and MRI-responder), true negatives (TN, true non-responder and MRI-non-responder), false positives (FP, true non-responder, but MRI-responder), and false negatives (FN, true responder,

but MRI-non-responder) were calculated and sensitivity and specificity were derived using the following formulas: Sensitivity = TP/(TP + FN); Specificity = TN/(TN + FP). The results are presented in the Table 3. The overall sensitivity was 94% and specificity – 75%. The lower specificity is driven by higher FN rate. This, in part, could be due to the limitations of both MRI and the neuropathology detection methodologies. MRI is sensitive to subject handling and hardware setup – motion artifacts and low signal-to-noise level could lead to higher T_2 values, producing false negative results. Neuropathology, on the other hand may not detect the lesion at the early or late stage of its developing or if inappropriate stain is used (e.g., silver stain can't reveal white matter damage) [163]. The latter event can produce apparent false positive result, which should be classified as a true positive.

In summary, T_2 MRI provides noninvasive readout of neurotoxicity in rats based on our studies using 10 neurotoxicants with varying degree of toxicity expression with preliminary general sensitivity of 94 percent and specificity 75 percent. It has a potential to become a non-invasive biomarker of neurotoxicity, which could be used in conjunction with current histopathological methods by guiding the location and time for brain sampling, or even as a surrogate endpoint.

Ira Krefting, MD, Jonathan Cohen, PhD, Shane Masters, PhD, Center for Drug Evaluation and Research/FDA, USA

Gadolinium-based contrast agents (GBCAs) used with magnetic resonance imaging (MRIs) improve disease visualization in multiple body organ systems and contribute to life saving diagnoses. GBCAs consist of a gadolinium (Gd) ion linked to an organic ligand either in a linear or macrocyclic structure; the linear agents have a greater propensity to de-chelate, meaning separation of the Gd from the ligand, than the macrocyclic GBCAs.

In the late 1990's some patients with severe renal failure undergoing MRIs with GBCAs were noted to develop a potentially fatal fibrotic skin and organ condition, when fully characterized termed, Nephrogenic Systemic Fibrosis (NSF). Free Gd from de-chelation is believed to play the primary role by initiating an immunologic cascade leading to NSF, and the search continues for possible co-factors. FDA contraindicated most of the marketed linear GBCAs in patients with severe renal failure and recommended screening renal function testing for patients with high-risk conditions associated with renal dysfunction. The practice community responded by largely transitioning to macrocyclic GBCAs and strictly adhering to the recommended dosing.

Around 2014, clinician researchers identified T1 MRI signal hyperintensity consistent with Gd in the dentate and associated nuclei in cranial MRIs performed without a GBCA but in patients who had previous GBCA exposure. Pre-clinical and human autopsy studies subsequently confirmed Gd in these brain regions and other organs. FDA received adverse event reports of pain and other symptoms following GBCA administration. An FDA advisory committee in 2017 could not identify any harm from Gd retention and could not attribute the patients' symptoms to GBCAs. The committee recommended a new warning to the GBCA prescribing information and a patient's medication guide to inform both providers and patients about potential adverse effects from Gd retention in vulnerable populations such as patients undergoing repetitive MRIs, the young, and pregnant patients. Studies continue the differential retention between linear and macrocyclic GBCAs, the molecular forms of retained Gd, the duration of retention and any possible pre-clinical or clinical effects.

FDA is awaiting completion of post-marketing studies investigating possible neurologic effects in neonatal and juvenile animals and in patients with chronic illnesses undergoing periodic, repetitive GBCA MRIs. Currently available published pre-clinical studies are summarized by Dr. Cohen and clinical studies are reviewed by Dr. Masters.

Since 2018, there have been several notable pre-clinical studies to identify effects following administration of GBCAs. These studies are significant for their behavioral and morphological findings but are limited by nonstandard study design, exemplified by differences in animal species and strain, gender, GBCA dosing, and recovery period, variables that make it difficult to assess positive findings in a clinical context.

Two more recent publications [165, 165], describe new morphological and behavioral findings, respectively, following administration of GBCAs that merit further evaluation of potential new safety signals and risk assessment. Radbruch demonstrated that linear and macrocyclic GBCAs reduced the intraepidermal nerve fiber density (IENFD) and increased the number of terminal axonal swellings (TAS) in the skin, 4-weeks post-dose. Effects on the IENFD and TAS/IENFD ratio were greater for linear versus macrocyclic agents and the authors attributed the histopathology findings to peripheral neuropathy of small unmyelinated fibers in the skin. Alkhunizi demonstrated Gd retention in the spinal cord and peripheral nerves (sciatic and trigeminal) following 20x repeat-dosing of GBCAs. Thermal and mechanical hyperalgesia (paw withdrawal from a heat source and by Von Frey test) were identified following repeat-dosing of gadodiamide (linear) but not gadoterate (macrocyclic), with dose-related findings for gadodiamide. However, interpretation of these new publications present challenges due to

differences in study design, for example animal species and strain, gender, GBCA dosing, route of administration, and recovery period. Moreover, there were differences in study endpoints that do not permit a direct comparison, for example. endpoints that were restricted to only morphological and IHC assessment or behavioral assessment.

Follow-up studies designed to evaluate potential peripheral findings should incorporate behavioral evaluation, detailed histopathology analysis, and inclusion of appropriate positive controls to aid in interpretation of these findings. Standardized assays and study designs are critical to evaluating risk based on nonclinical findings and to advance regulatory science.

Based on the extensive clinical experience with GBCAs, adverse events related to Gd retention are expected to be relatively uncommon, so large sample sizes are needed to appropriately power most study designs. One of the larger studies to search for clinical effects of Gd retention was performed by McDonald [167], who analyzed a prospective cohort of 4261 cognitively normal subjects enrolled in the Mayo Clinic Study on Aging, 1092 of whom had received at least one dose of a GBCA. After adjusting for potential covariates, no association between GBCA exposure and cognitive decline, neuropsychological performance, or motor performance was found. In addition, no effect of the number of doses of GBCA on these factors was observed.

For studies that do find potential adverse events related to Gd retention, accounting for potential confounding is often difficult. For example, a study in patients with multiple sclerosis, who may be at greater risk due to neuroinflammation and exposure to multiple doses of GBCA, found a correlation between cognitive function and T1 and T2 relaxation rates, which are markers of brain Gd retention [168]. Specifically, these markers of Gd retention in the thalamus and dentate nucleus were associated with decreased information processing speed, while in the thalamus and caudate nucleus they were associated with decreased verbal fluency. The authors do note the possibility that despite attempts to control for multiple sclerosis severity, the disease status may confound these results.

A prospective study investigated the incidence of symptoms similar to those reported as adverse events using a directed survey of 607 patients receiving gadodiamide or gadoterate with MRI compared to 481 patients receiving MRI without contrast[169]. Mental confusion and fatigue, but not other prespecified symptoms, were found to occur more often in patients receiving GBCA during the 24-hour study period. One patient reported symptoms lasting up to 2 months after the MRI, however the symptoms of all other patients resolved within 24 hours. The results may be

confounded by lack of blinding and by differences in indications for MRI between the GBCA and non-contrast groups.

Neuro-cognitive adverse effects from Gd administration remain in question. An FDA post marketing requirement study will follow patients for approximately 5 years with functional testing as they undergo routine MRIs with GBCAs for chronic clinical conditions such as liver disease or prostate cancer.

John Waterton, PhD, University of Manchester and Bioxydyn Ltd, United Kingdom

Imaging biomarkers are used in safety assessment all along the drug development pathway. In animal studies, they may be used to avoid dangerous candidates ever getting into man, to prioritize the safest compounds, or to find dose-limiting toxicities and define the therapeutic margin. In early drug development, a safety signal might be a good reason to stop the project. In late-stage development, imaging biomarkers can be used to show reversibility, or in labelling either for contraindications or for monitoring.

There are different challenges in validating imaging biomarkers, compared to the more familiar biospecimen biomarkers. A biospecimen biomarker usually requires extraction of something from the patient, such as a tissue or biofluid, which is taken to a lab for analysis using a dedicated in vitro diagnostic device. Imaging biomarkers, however, use whatever scanner happens to be available in the particular hospital that the patient is attending. These imaging devices are not designed or maintained for measuring imaging biomarkers, and different makes and models from different vendors may provide measurements that are not directly comparable. In vitro devices are often designed and operated in central labs, by staff trained for the specific assay, with excellent quality control, specifically for the purposes of analytical biochemistry. There is usually a defined analyte, and validation projects can spike with authentic material, assess limits of detection [170]. In contrast, imaging biomarkers seldom assess anything that could be recognized or defined as an analyte. Indeed, imaging vendors are often not very interested in quantitation: their business is to make beautiful pictures that help radiologists make diagnostic decisions. Sometimes, when the manufacturer improves the picture quality, there are unexpected effects on quantitation: for example, in the introduction of parallel imaging into MRI, which changed the noise characteristics of the underlying scans, and potentially the error propagation into the biomarker. Consequently, imaging does need a different validation roadmap, and specific imaging biomarker roadmaps have been devised [171]. The key insight is that, while

biospecimen biomarkers usually start by validating the assay and locking it down before it is used in substantive studies, for imaging biomarkers the process is iterative: partial assay validation, then early use to gather initial biological information; small clinical trials to explore the potential of the assay; strengthen the assay validation platform to support multi-center multi-vendor use; larger trials; and ongoing revalidation to keep up with the evolution of the installed base of scanners. For use in clinical trials, we need enough confidence in the imaging to answer trial questions, and multi-center trials need more evidence than single-center trials. Continual iterative validation is needed for healthcare use. So, for investigators who are used to validating biospecimen biomarkers, the approach to imaging may feel rather informal and unstructured.

Furthermore, imaging biomarkers of harm are even harder to validate than the more conventional imaging biomarkers of benefit. A trial with possible benefit to participants shouldn't usually have major ethical concerns. On the other hand, a trial which deliberately sets out to harm the participants so that the toxic effect can be compared with the imaging biomarker would likely be ethically unacceptable. For a rare occurrence of harm which is unplanned, scheduling the imaging visits to coincide with unplanned harm is much more challenging than scheduling a regular imaging follow-up after planned treatment for benefit. An important statistical point is that imaging biomarkers of benefit often compare pre- versus post-, so the relevant statistical metric is repeatability: same subject, same scanner, same observer. On the other hand, imaging biomarkers of harm often a lack a pretreatment scan (because we never intend to harm the subject), and the post-treatment imaging biomarker must be compared with a normal population range, so the relevant statistical metric is reproducibility: multiple international centers, multiple equipment, multiple investigators. Hence imaging biomarkers of harm often require more extensive characterization of assay variability than imaging biomarkers of benefit.

Another consideration is that in studies of biomarkers of benefit, physicians are often very enthusiastic to recruit, because they want to demonstrate benefit to their patients, but may be understandably less enthusiastic about measuring any harms caused to their patients. Also, drug developers and pharma companies may be rather interested in novel biomarkers of benefit, even if not fully validated, because of the potential for different insights into the value of the investigational drug. On the other hand, drug developers are understandably extremely reluctant to use unvalidated biomarkers of harm, which risk measuring something which is uninterpretable yet alarming to patients, physicians, and regulators.

So how can imaging biomarkers of safety be developed? One very powerful approach is public-private partnerships. In Europe, the Innovative Medicines Initiative has invested over €5 billion

into consortia involving pharma, maybe imaging or health tech companies, small businesses, academia, and other stakeholders. TRISTAN *(https://www.imi-tristan.eu)* is an IMI project validating imaging biomarkers of drug safety. It focusses on three areas of toxicologic interest: drug-induced perturbation of liver transporter fluxes; drug-induced interstitial lung disease, and the harmful maldistribution of biologic drugs. The consortium's aims are to verify their imaging biomarkers against the underlying (ground truth) biology; to standardize assays; to translate them between animals and humans; to ensure they're reliable multi-center, and not just for a single study in a single expert center; and ultimately, to make assays available both to commercial clients, such as pharma companies, and also to academics as published methods that they can use in their own research.

The liver is the most common organ associated with drug withdrawals, and frequently liver injury (DILI) is not predicted from preclinical studies. Drug-induced inhibition of transporter fluxes can cause both DILI and drug-drug interactions. MR contrast agents, particularly gadoxetate, and PET tracers, can be used develop in vivo assays for relevant transporters, such as OATP and BSEP [130]. The great advantage of the imaging approach is that these assays employ time-resolved measurements of concentrations in the blood, hepatocyte, bile, renal, and enteric compartments, making it easier to parameterize the models and to make accurate and highly precise measurements of the fluxes through transporters and perturbations by investigational substances. Gadoxetate is used in liver patients, approved by regulatory authorities in most jurisdictions and routinely used by radiologists. Uptake of gadoxetate into the liver, and then the excretion into bile and into the urine and the enteric system are perturbed, particularly into bile, both by investigational drugs [133], also with more established drugs [132], such as rifampicin [134].

Drug-induced interstitial lung disease (DIILD) is a recognized subtype of interstitial lung disease that has been a somewhat neglected area in drug safety assessment. At least 27 drugs have been associated with DIILD [151]. It is serious enough to be included in the prescribing information, and in three cases, as a boxed warning [152]. It can involve immuno-oncology drugs as well as anti-infectives and older cancer drugs, anti-inflammatories, and so on. While chest radiography is sometimes advised in the prescribing information [152] there is little use of modern imaging modalities or of cut-offs or decision trees. TRISTAN employs modern modalities; CT, 129Xe MRI, proton MRI (including oxygen-enhanced and gadolinium-enhanced), and some PET biomarkers. Our philosophy, because we do not want to deliberately harm patients, has been to recruit patients with suspected DIILD identified as part of their standard care so the intervention is the withdrawal of the drug, and the hypothesis is change in imaging biomarker following withdrawal.

TRISTAN aims to validate new imaging biomarkers in these important areas. The consortium is working to prove its imaging assays are reproducible between different labs and different vendors, using clinical studies which physicians are interested in recruiting into, and ultimately to give drug developers and regulators the confidence that these measurements are worthwhile approaches to measuring drug-induced harm, reversibility or lack of harm.

Timothy J. McCarthy, PhD, Pfizer, USA

The role of imaging as a key biomarker in drug development has evolved significantly over the years [172]. In part this has been enabled by the concept of three pillars [173], which provide context for how an imaging modality can provide answers to questions around drug distribution, target engagement or downstream pharmacology. While many examples exist that demonstrate evidence to support proof of mechanism, there are less examples of imaging as a biomarker of safety. One of the key reasons for this is that, by definition, assessments need to be carried out across large populations and, unlike proof of mechanism studies, cannot be demonstrated in small populations of healthy volunteers. As an example, consider the use of PET to map out the occupancy exposure relationship of a $GABA_A$ modulator [174], which can be achieved in healthy volunteers using a limited number of participants at a single site.

Using imaging in the context of a safety biomarker requires access to the given technology at all sites enrolled in the study. This immediately precludes the use of cutting-edge approaches, such as the PET example cited above and forces consideration of methods that are widely available in a community setting in which the required methods do not deviate far from standard radiological workflows. In addition, analysis and interpretation of results needs to be rapid, in order to provide immediate feedback to the site for the well-being of the participant. One common feature of imaging when used as a biomarker of safety or efficacy is that the concept of interest and context of use needs to be carefully qualified and validated. This framework is central to the practices detailed by the FDA-NIH Biomarker working group at the BEST (Biomarkers, EndpointS, and other Tools) Resource (https://www.ncbi.nlm.nih.gov/books/NBK326791).

In order to illustrate the challenges of deploying imaging as a biomarker of safety, consider the risks of cardiotoxicity in cancer therapy [175]. This includes any structural or functional heart injury related to cancer treatment. Injury most commonly involves the myocardium leading to heart failure (HF) but can also include the pericardium, valves or coronary arteries.

Cardiotoxicity is defined as a decrease in left ventricular ejection fraction (LVEF) of ≥5% to <55% in the presence of symptoms of heart failure or an asymptomatic decrease in LVEF by ≥10% to less than 55%. There are several methods to measure ejection fraction, these include echocardiography, cardiac MRI, and radionuclide (MUGA) scans. Each has its benefits and limitations, but considering the considerations laid out earlier, echocardiography has significant advantages in terms of technical simplicity, availability and cost of deployment. Guidelines from the American Society of Echocardiography and the European Association of Cardiovascular Imaging have been published [156] to identify best practices for assessing LVEF in the context of monitoring cancer therapy. It would be remiss to ignore efforts to identify circulating biomarkers of cardiotoxicity and significant research has been conducted in this area [176]. Brain natriuretic peptide (BNP), N- terminal proBNP and Troponin have emerged as promising biomarkers. In the case of Troponin-1, studies [177] have suggested that assessment of levels during trastuzumab therapy may allow for identification of patients at risk of cardiotoxicity as well as those who, despite HF therapy, will not recover from cardiac dysfunction. Looking to the future it is likely that a combination of circulating biomarkers in combination with readily available imaging techniques will provide an excellent means of monitoring patients on novel therapies or approved combination therapies, post-approval.

Another interesting example of the role of imaging in safety can be found in the field of osteoarthritis and the development of novel therapies targeting the nerve growth factor receptor. Work by Roemer and colleagues [141] has led to the development of an imaging atlas using both radiography and MRI to guide enrollment eligibility and on-study monitoring of participants who are enrolled in trials of these agents.

To summarize, the application of imaging as a tool to monitor the safety of novel and approved therapies has a lot of potential, but in comparison to the more traditional use of imaging as a tool for mechanistic studies, there are some significant differences and limitations that must be considered. This fundamentally means that the modalities and techniques used should be cost-effective and widely available in the community setting with familiar acquisition and analysis techniques, ideally supported by established guidelines from leading academic societies. Turn-around-time for processing and interpretation needs to be rapid in order to ensure delivery of the best care to the patient. Another consideration is the interplay between blood-based biomarkers for safety and imaging; how together they may provide more sensitivity, or an opportunity to use early changes in the circulating biomarker to prompt the need for imaging follow-up. The use of combination therapies is increasing across several diseases and long-term follow-up of patients will be key for these treatments in a post-approval setting. Finally, the application [178] of machine

learning and artificial intelligence to radiology, image analysis and disease prognosis will clearly play an important role in this challenge moving forward.

Yan Liu, MD, PhD, Median Technologies, France

Cancer is the second leading cause of death globally *(https://www.who.int/news-room/fact-sheets/detail/cancer)*. Within oncology, medical imaging is key to the patient journey. Billions of radiological images are generated per year to support diagnosis and monitoring *(AI in medical imaging forecast. 2020. https://www.signifyresearch.net/medical-imaging/ai-medical-imaging-top-2-billion-2023/)*.

Those images also offer quantitative, non-invasive, and highly standardized data, and would be a great source for data mining, to establish a new, affordable mode to find cancer cures.

Today, radiologists make diagnoses based on medical images, clinical information, and their experience. The quality of medical imaging has dramatically improved via innovations in radiological equipment. Systematic investigation of computer-aided diagnosis (CAD) began in 1980s, and today, CAD tools are a routine part of detecting lesions in breast cancer and lung cancer, as well as other screening procedures. This is just the starting point; we believe next generation artificial intelligence (AI) will push the boundary even further, thanks to the huge amount of data archived in standardized systems (such as Picture Archiving and Communication Systems) and enhanced technology and computing powers.

The clinical value to medical imaging of today's AI tools is in disease detection, quantification, and characterization, with the main aim to reduce the radiologist's workload and improve diagnostic accuracy. Next generation AI will be revolutionary, transforming medical images into mineable high-dimensional data. This predictive value could become a powerful tool for patient stratification and treatment management.

When applied to standard medical imaging, current AI technologies have the ability to model biological complexity and classify data patterns by assessing multidimensional clinical and biological data. Considering oncology, accurate prediction models are the central component in developing imaging biomarker panels that could provide insights on host environment, tumor and tissue heterogeneity, and tumor microenvironment. AI-based imaging solutions can be used for molecular subtyping, risk stratification, and prediction of treatment response and thus hold the promise of delivering precision treatment in both drug development and patient care.

Median Technologies is developing a new-generation AI-based imaging platform, iBiopsy®

(Imaging BIOmarker Phenomics SYstem). iBiopsy® is a platform that can extract comprehensive digital imaging signatures from the whole organ, using advanced mathematical learning models for prediction, prognosis, and diagnosis.

In contrast to traditional radiomics [179], iBiopsy®'s technology is based on whole-organ analysis, therefore avoiding variability and error that can occur when only regions of interest (ROIs) are considered in images. Whole-organ analysis can also apply to multiple organs, enabling disease- and stage-agnostic prediction. Inputs include medical images and multimodal data (e.g., genomics, patient outcome) to develop AI fit-for-purpose tools. Eventually the classification should allow patient stratification.

iBiopsy®'s machine learning algorithms was applied to Magnetic Resonance Elastography (MRE) images to predict liver fibrosis scores in nonalcoholic steatohepatitis (NASH) patients and discriminate between early- and advanced- liver fibrosis. In its early stages, NASH remains reversible via diet and lifestyle changes, whereas, in its end stages, liver transplantation may be a patient's only treatment option. As fibrosis grade is key in NASH prognosis, there is clinical interest in establishing reliable and non-invasive tests to distinguish accurately patients with early versus advanced fibrosis. Liver biopsy, though used for diagnostic purposes, is not attractive as the first indication for a number of reasons.

The NASH clinical research network divides fibrosis into: F0-1: absent or mild fibrosis; F2: significant fibrosis; F3: bridging fibrosis; and F4: cirrhosis. Of these, only cirrhosis can be detected by radiologists on computerized tomography (CT) or magnetic resonance images (MRI). Using magnetic resonance elastography (MRE), which shows stiffness of liver tissue, helps to indirectly measure liver fibrosis. In a first study, using MRE images, we compared performance in prediction of histological fibrosis scores between imaging features learned by iBiopsy®'s Convolutional Neural Network (CNN) and those hand-crafted by experienced radiologists; we found iBiopsy® to better predict advanced fibrosis (Figure 33 [180]]) .

Liver fibrosis is also recognized as a good predictor for hepatocellular carcinoma (HCC) recurrence [181]. In another study, we applied iBiopsy®'s algorithms to CT images from a cohort of 160 HCC patients. The iBiopsy® low-risk score had a significant recurrence-free survival benefit compared to the iBiopsy® high-risk score, with a hazard ratio of 4.1; while the histology fibrosis score gave a hazard ratio of 6.6 ; [182]]. The advantage of learning models is the possibility to add other features to improve prediction. A combined model (using both histology and iBiopsy®) demonstrated excellent prediction of outcome for patients when both histology and AI confirmed low or high risk (Figure 34).

Immunotherapy is changing the landscape of anti-cancer therapy. However, more than 50% of

patients do not respond to immune checkpoint inhibitors (ICIs) and response is even worse in some indications; the objective response rate in patients with primary liver cancers is about 20%, for example. Imaging biomarkers allow better understanding of tumor-host interactions, e.g., CD8 infiltration. Immune-inflamed tumors are characterized by dense, functional CD8 infiltration, which provides a higher chance to respond to ICIs.

A study examined the performance of iBiopsy®'s AI-based imaging solution to predict CD8 infiltration in HCC patients. Compared to traditional radiomics features, iBiopsy® deep-learned features had lower mean average error. In addition, with iBiopsy®, no lesion segmentation was required. Similarly, when a deep CNN with attention mechanism was applied for CD8 high or low prediction, it better predicted the immune microenvironment than traditional radiomics, with an area under the curve of 0.93. Thus, AI approaches could help us better understand tumor microenvironment when tissue samples are not available, with the potential to guide therapy.

There are benefits that AI could bring, but in parallel, one must realize the potential risks of implementing AI-based diagnosis for patients. Key considerations are:

- How to deal with AI-generated false positives or negatives that would have been diagnosed by a clinician?

- Can AI developed on virtual data sets be translated into real-world data?

- Are AI tools developed in specific population for specific use at risk of inappropriate "off-label" use?

In this regard, the Food and Drug Administration (FDA) has published a discussion paper on the regulatory framework for modifications to AI/machine learning-based software as a medical device (SaMD) *(AI/machine learning based software as a medical device (SaMD). 2020. https://www.fda.gov/media/122535/download)*

. To date, the FDA has approved several AI-based SaMDs, whose algorithms were typically 'locked' prior to marketing. However, since algorithms can adapt over time, we need a practical regulatory approach that allows such devices to improve, while still providing effective safeguards.

We support fully the FDA's suggestions. An approach is needed which enables early-stage development with a view to life-cycle enhancement through AI, appreciates adaptive learning systems, and bases regulatory oversight on inputs/outputs rather than process. To allow good products to become available: risk categorization should not limit the application of continuous

learning systems, Good Machine Learning Practices should not burden developers, and active dialogue with regulators about real-world continuous learning is needed.

Track E: Microbiome, Reinhilde Schoonjans, PhD, and George E. N. Kass, PhD, European Food Safety Authority, EU

The microbiome found in the food chain and the human gut have attracted considerable attention from many research communities and commercial organizations, with the focus on a sustainable and healthy future. While many scientists view this topic from the perspective of applications, investment and commercial opportunities, an urgent question is how risk assessors and regulatory scientists can engage to ensure the safety of the consumer is not compromised.

The microbiome refers collectively to communities of microorganisms and their genomes in a defined environment. Microbiomes are composed of many life forms including bacteria, archaea, viruses, or eukaryotic microorganisms (such as protists, fungi, and algae). Because of their ubiquitous presence, microbiomes occupy a central position in the "One Health" framework, which approaches human, animal, and plant health from a new integrated perspective.

The microbial ecosystem most explored is the microbiome of the human gut. This microbial community interacts with the host and the intestinal mucosa. There is mounting evidence for a role of the gut microbiome in several enteric and systemic disorders in humans. Therefore, interactions between the microbiome and an environmental chemical modulator might influence the health of the host through direct chemical-induced changes in the microbiome composition. The capacity of chemical modulators to induce microbiome changes in animals has been demonstrated with a variety of pesticides, metals, artificial sweeteners, and drugs.

In the absence of explicit legal requirements to account for effects on the microbiome in risk assessment, there is no guidance or methodology in place to account systematically for potential effects of regulated substances on the microbiome or effects mediated by the microbiome on human, animal, or plant health. Session E of the GSRS2020 offered a direct opportunity to reflect on the current knowledge of the human gut microbiome, the unknowns and the direction regulatory science could take in terms of methodologies, data generation and appraisal of evidence. The purpose of this session was, therefore, to discuss with regulatory scientists around the world the risk assessment questions to be asked and the appropriate

starting points for meaningful action in this field, to exchange on ongoing efforts and to contribute to capacity building in this area.

Martin Iain Bahl, PhD, Technical University, Denmark.

The issue of incorporating effects on the human microbiome in chemical risk assessment is described. Considering the human microbiome, it is important to appreciate that the microbial communities associated with humans have co-evolved with us through millions of years of evolution. It is known with certainty that the commensal bacteria colonizing the gut both play an important role in metabolism and profoundly influence human health in many ways. This means that for risk assessment we might consider human beings as holobiont organisms comprising both our own human genome and that of the hundreds of different symbiotic bacterial species associated with it [183, 184]. Potential negative effects on the host-associated microbial community should also be addressed when considering overall risk assessment of environmental pollutants. Focusing more specifically on oral exposure to pesticides, many different substances, including herbicides, insecticides and fungicides, have been shown to affect the microbiota composition and or activity when tested under controlled conditions *in vitro* or in laboratory animal exposure trials. An example of this is exposure to glyphosate, which is the active ingredient in several commercially available herbicides including Roundup®. For this particular substance there is a clear mode-of-action attributed to inhibition of bacterial growth, since many bacteria encode a homologous target pathway for aromatic amino acid synthesis (phenylalanine, tyrosine, and tryptophan) as found in plants – namely the Shikimate pathway [185]. Several *in vitro* trials have indeed shown that bacteria colonizing the gut are also targets for this pesticide [186], and furthermore, some bacterial species seem to be more sensitive than others when exposed to glyphosate [187], which could affect the bacterial community composition in the gut. An animal exposure trial in rats conducted by me and colleagues [188], however suggested that endogenous levels of aromatic amino acids in the gut environment derived from dietary components might alleviate the effects on bacterial growth following oral exposure. This finding is further supported by a study showing a higher degree of incomplete Shikimate pathways in the genomes of host-associated bacteria compared with free-living bacterial species [189]. Despite these observations specifically for glyphosate, pointing towards limited effects on gut bacterial communities, it seems clear that some environmental pollutants do have the capacity to affect bacterial communities in the gut when challenged with sufficiently high concentrations. However, many natural compounds and dietary choices [190] also affect bacterial composition, so

to base risk assessment solely on any measurable change of the gut microbiota does not seem useful. It is suggested to work towards developing tools, markers and endpoints to assess "microbiota disruption", which could be defined as induced changes in microbiota composition and/or activity that are causative to a detrimental health effect. Within the microbiome research field, the use of germ-free animal models to demonstrate causality has become the gold standard. Such studies may in the future help pinpoint which changes in microbiota composition or activity should be considered detrimental and thus included as markers in risk-assessment exposures trials for environmental pollutants.

Joseph V. Rodricks, PhD, Ramboll USA

It is reasonable to postulate that the microbiome can play a role in the development of chemical toxicity. It is well established that alterations in the microbiome can lead to adverse health outcomes of several kinds. It is also well established that exposures to some chemicals can alter the microbiome. What remains to be determined is whether chemically induced microbiome perturbations of specific kinds can induce adverse health outcomes (toxicities) separate and apart from those resulting from well-known toxicity mechanisms.

It is also well established that uptake and metabolism of some chemicals can be altered or modulated in several ways by the microbiome. The consequences for a chemical's toxicity profile of these microbiome pathways are incompletely understood (Figure 35).

Although research on these topics is moving ahead on several fronts, it is not now possible to evaluate the importance of these possible mechanisms of microbiome influenced toxicities. It is far from clear that current methods for studying chemical toxicities – animal and other experimental studies and observational epidemiology studies – can provide readily identifiable evidence on microbiome influences. If those influences are significant, then risk assessments based on such studies may not provide adequate characterizations of human risk [191].

What is the existing evidence that particular chemical exposures or other interventions -- dietary interventions, for example -- can affect the assembly, the maturation, or in germ-free mice? (Figure 36 and references [192, 193]).

In one study, mice were administered saccharin at doses equivalent to the usual intake and glucose intolerance was observed. Fecal microbiomes from unexposed animals were treated

with saccharin in vitro and used to colonize germ-free mice. Glucose intolerance was observed in these mice, suggesting a direct causal relationship between chemically induced changes in the microbiome and host response [194]. In recent years there has been much study of the possible effects of non-nutritive sweeteners on the gut microbiome, with possible metabolic effects in the host. The suggestive evidence from some of these studies does focus attention on the possible significance of microbiome perturbations and the need for research in this area of food safety.

One significant consequence of finding a significant role for the microbiome in the development of toxicity concerns the important matter of variability in response. The composition, gene content, and functional characteristics of the microbiome vary considerably from gestation to death, and by gender and race, pregnancy status, diet, and geography, and by body sites (and niches within those sites). By whichever mechanistic pathway the microbiome influences toxicity outcomes, greater variability in response is expected than will be seen in the absence of such an influence. Thus, arises the question of whether current uncertainty factors used to account for variability are adequate. Is it possible, for example, that often-observed inconsistencies in findings from epidemiology studies are due in part to microbiome-influenced differences in the populations studied [191].

Considerable research is necessary to clear the way to an understanding of the roles of the microbiome in chemical toxicities, and the implications of these roles for human risk assessment. It is also clear that that understanding is necessary to ensure fully adequate food safety decisions.

Key Points

There are good reasons to suspect that some interactions between chemicals, including food additives, and the gut microbiome can result in adverse health effects that are specific to those interactions.

Microbiome mediated effects cannot be recognized in standard animal toxicity studies or even in epidemiology studies in selected populations. Specific study designs are necessary to uncover such effects.

There is yet only indirect evidence that such effects are important in the development of chemical-related adverse health effects, but that evidence strongly suggests the need for targeted research on the subject.

Several issues relating to dose-risk relationships, inter- and intraspecies extrapolations, and exposure assessment need to be incorporated into research on chemical-microbiome interactions. These subjects have yet to receive significant attention in microbiome related research

It is also likely that targeted microbiome-related research will reveal valuable information about chemical risk, particularly regarding variabilities in response.

Carmen Pelaez, PhD, Spanish Council for Scientific Research, Spain

Diet is one of the major factors highly influential in shaping the gut microbiota. Westernized long-term diets and some additives and chemicals have been increasingly associated with chronic non-transmissible diseases. However, there is no methodology in place to assess the impact of food and chemicals on the gut microbiome and its potential consequences in human health.

In this overview, general considerations concerning the effect of food and chemicals on the gut microbiome is discussed, as well as a number of model systems that can be successfully used in a tiered approach for the risk assessment of food/chemicals including the gut microbiome

The intestinal tract is inhabited by a large and diverse population of microorganisms (and their collective genomes), termed the gut microbiome. This microbial population connects the human internal physiology with the external environment and display a vast array of functions that influence metabolic, immune, cognitive and defense systems, consequently having an impact on human health [195].

The gut microbiome is largely influenced by major environmental factors such as diet, lifestyle including physical exercise and exposure to antibiotics. But, also age, geographical location and habits linked to traditional or Westernized long-term diets play a very relevant role in the complexity of the gut microbiota and in some cases have been associated with the emerging increase of chronic non-transmissible diseases, such as obesity and associated metabolic diseases, inflammatory bowel diseases, colorectal cancer, and allergies, among others [193]. Also, the increase in scientific evidence over the last years about the potential interaction between some food additives/chemicals and the human gut microbiome [197, 198], reinforces the need to revise current risk assessment of food and chemicals by incorporating their effects on the gut microbiome.

So far, there is no guidance or methodology in place to systematically assess the impact of food and chemicals on the gut microbiome. Therefore, there is a need for validated methodologies that may be used for this purpose and further on, to test the hypothesis of causality for potential microbial negative effects in a dose-response manner.

A standard protocol for risk assessment includes some major areas of action. Among others, these areas are: 1) to identify potential hazards, 2) to define the core areas of interest and the outcomes that relates to toxicity and 3) to characterize potential hazards using models.

The first area that need consideration during the assessment is the toxicokinetics of the study compound that includes its potential metabolization but the gut microbiota. This follows a conventional risk assessment protocol that includes *in vitro* bioaccessibility (static and dynamic GIT models including microbiota), intestinal permeability (USSING Chamber, mucosal sheets, everted sacs, epithelial cells) and *in vivo* bioavailability (*in situ* perfusion, oral bioavailability). The second area involved is the potential impact on microbial structure and loss of diversity. This can have a profound effect on the gut defense mechanisms leading to an increased risk of infections. Third major area is whether the compound under study is up or down regulating the microbial metabolic gene expression and unbalancing the short-chain fatty acids production or any other key metabolites such as neurotransmitters, that can trigger metabolic diseases or cognitive disfunctions.

And the fourth area to study is the microbial-host interaction and whether there is any loss of mucosal integrity that can impair gut defense mechanisms and trigger inflammatory diseases and allergies.

To characterize potential hazards, analytical tools to assess the effects on the gut microbiome, are needed. The microbiome community structure can be profile by using metagenomic whole genome shotgun sequencing that allows microbiome species-strain-level resolution and functional gene assignment. Furthermore, multi-omic approaches including meta transcriptomics and metabolomics can be combined to understand the dynamics of transcriptional regulation and metabolite production. In silico analysis is also a powerful tool to reconstruct metabolic networks at genome scale and predict metabolites produced by the gut microbiome [199].

Despite the enormous potential of the technologies mentioned above to point out associations between chemicals and human diseases, there is still a need for experimental work. *In vitro* and *ex*

vivo systems allow for a better control of conditions than *in vivo* models. Flow dynamic simulators mimicking the gastrointestinal tract are highly reliable for microbial-chemical interaction studies, but they typically lack human cells, although this can be overcome by introducing mucin-covered microcosms in the reactors [200-206]. An example of a flow dynamic simulator is the BFBL simulator developed in Spain which includes a small intestine and three reactors simulating the ascending, transversal, and descendant large intestine. Stabilization of composition and metabolism of fecal microbiota takes place over 14 days, after which the system is fed with the compound under study [207].

Recent developments on polarized epithelial cell lines grown on 3D scaffolds or even more sophisticated microfluidic devices that allow cyclic peristalsis –like motions and fluid flow, can facilitate co-culture of human intestinal cells with commensal microbiota for extended times and can be used to study barrier function and host-microbiome interactions [208]. They have a potentially moderate physiological relevance, but in some instances, they still need validation for complex microbial communities.

Beyond *in vitro* experiments, further *in vivo* models to validate the results are needed. The most promising models to study causation of diseases by a compromised microbiome is the germ-free murine model (gnobiotic) as well as rodents with depleted microbiotas, that are transplanted with human microbiota (so called Humanized Mice) [209]. The main limitation of these models is the lack of co-evolution between the host and the microbes of the donor that leads to differences in the microbiome-host interaction between the donor and the recipient species.

Conventional risk assessment for food and chemicals following OECD Guidelines, includes a tiered approach where results of studies at higher tiers will supersede results at lower tiers. When the risk assessment takes into consideration the impact on the gut microbiome, a similar tiered approach could be proposed (Figure 37).

Tier 1 based on initial *in vitro* studies can be designed in order to obtain reliable data on the effect of the food/chemical on the gut microbial structure and function, as well as the dose/response relationships. Should the obtained data indicate preliminary links between the substance under study and the perturbed gut microbiota, proceeding to Tier 2 with the use of *in vivo* animal models and additional human data, when available, will provide useful information to establish causality on the final health outcomes.

In summation, models to assess the potential risk of food and chemicals including the gut

microbiome are increasingly emerging, but some of them may still need certain adaptations to be applicable in standardized protocols. Combined multi omic techniques will be extensively used to address the complexity of the gut microbial community structure and function in terms of variability and temporal dynamics. Overall, more research   work on risk assessment to go beyond simple associations until final causation, is still needed.

Conclusions:

It is evident from the previous overviews that emerging technologies will play a major role in regulatory science in the future. One could argue that there has been an evolution of use and incorporation of new approaches from the very beginning of the safety assessment process. As the pace of development of novel approaches escalates, it is evident that assessment of the readiness for these new approaches to be incorporated into the assessment process is necessary. By examining the arears of Artificial Intelligence and Machine Learning; Omics, Biomarkers, and Precision Medicine; Microphysiological Systems and Stem Cells; Bioimaging and the Microbiome, clear examples as to how to assess the reproducibility, reliability and robustness of these new technologies have been revealed. In a group movement, there is a call for product developers, regulators, and academic researchers to work together to develop strategies to verify the utility of these novel approaches to predict impact on human health. When that occurs, those new technologies that reliably and routinely provide valid data for safety assessments as compared to existing procedures and reflecting human health effects will be incorporated into acceptable testing regimes.

Authors' Contributions

Drs. Arnd Hoeveler, Marta Hugas, Anil Patri, Danilo Tagle and Neil Vary all contributed to the planning/selection of the co-authors as well as the review of the manuscript. The other co-authors contributed written segments as well as review of the manuscript.

Disclaimer

The opinions expressed in this paper are those of the author, and do not necessarily represent the views of the US Food and Drug Administration or the US Government. The views expressed in this article are the personal views of the authors and may not be understood or quoted as being

made on behalf of or reflecting the position of the agencies or organizations with which the authors are affiliated. No endorsement or recommendation is inferred from the mention of brand names or descriptions of approaches. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization or other government agency.

Figure Legends:

Figure. 1: Definition of machine learning and deep learning.

Figure 2: Steps in product development.

Figure 3: History of the Sentinel Initiative.

Figure 4: Sentinel System Innovation Strategies.

Figure 5: Four-pronged approach to Sentinel Innovation Strategy Implementation.

Figure 6: Sample cloud tag for 'severe adverse events', the actual cloud tags contained more than 14,000 terms.

Figure 7: Description of main steps of the performed systematic review with corresponding automation tools and obtained numbers of publications.

Figure 8: Predictive, personalized, participatory, and precision health (i.e., pHealth) enabled by the advancement of emerging biomedical technologies. Multi-modal biomedical data of the same patient are combined by AI-models to provide a comprehensive evaluation of the patient.

Figure 9: An exemplary pipeline for AI-enabled pHealth model using RNA-seq data. A. Raw -omics data are collected with high-throughput sequencing techniques. B. Multi-omics features are extracted with various bioinformatics pipelines. C. AI-based predictive models are utilized.

Figure 10: Precision was uniformly high across the ctDNA assays.

Figure 11: The PCA indicates the variation in the experiments did not alter the outcome to any large extent. After PLS-discriminate analysis, we were able to cluster the samples based on how we processed them. The ones that showed the largest effect was the blood left at room temperature for six hours or the plasma which was left at room temperature for 24 hours.

Figure 12: The palmitoyl carnitine peak appears before ALT peak as observed in rodents and humans.

Figure 13: Approximately 3,800 peaks differentiated high opium users from non-opium users, and approximately 712 peaks differentiated high opium users diagnoses as OUD positive from high opium users diagnosed as OUD negative.

Figure 14: An outline of the role and impact of genomics, proteomics, and metabolomics in assessing the safety and evaluating the provenance of foods, drugs, and cosmetics for regulatory science.

Figure 15: A typical workflow for *in silico* metabolomics and the utility of BioTransformer in identifying chemical "dark matter".

Figure 16: Initiatives that have contributed to metabolomics scientific reproducibility and standardization.

Figure 17: AMED-MPS Project in Japan. Researchers from the National Institute of Advanced Industrial Science and Technology (AIST) and the National Institute of Health Sciences (NIHS) coordinate the Device Manufacturing program, Standardization program of MPSs developed by universities and other academic institutions in Cell supply and Model development program. Researchers from pharmaceutical companies participate in Standardization program. Central Research Center provides a place for research and development.

Figure 18: Hepatic Sinusoid and regional function. Hepatocytes within the lobule shows different functions in region dependent manner.

Figure 19: Standard procedure of the acceptance of new test method as a test guideline in OECD.

Figure 20: Overview of the European Roadmap for OoC Development.

Figure 21: Value chain and roadblocks from an MPS invention towards a validated MPS-based assay benefiting patients.

Figure 22: Constructive interdependent qualification processes for MPS-based context-of-use assays at TissUse. Equipment qualification includes Installation qualification (IQ), Operational qualification (OQ) and performance qualification (PQ) at end-user sites after technology transfer.

Figure 23: Readiness levels defined at TissUse to easy customer communication, collaboration, and technology transfer management.

Figure 24: MSC aggregates in chondrogenic condition. (A) representative images of two MSC lines (MSC-A and MSC-B) at early passage (passage 2 or 3) at day 7 (D7) and day 21 (D21); (B) MSCs at late passage (passage 5) at D7 and D21; (C, D) histology images that were analyzed to confirm the deposition of cartilage-associated extracellular matrices (GAG shown in blue and collagen shown in red). The histology data show that the cells lines recovered in size deposited higher amounts of cartilage-associated extracellular matrices compared to the cell lines did not show the recovery (D).

Figure 25: Regulatory requirements and measures.

Figure 26: BBB chip with iPSC-derived brain microvascular endothelial cells (BMECs) and neural cells.

Figure 27: Motor neurons from iPSCs derived from patients with ALS. These have been seeded onto chips along with BMECs from the same iPSC line.

Figure 28: A tiered approach to regulatory science – starting with 2D-based screening, but then proceeding with more biologically-relevant 3D human models.

Figure 29: Tiered workflow for establishing the reproducibility and context of use for MPS at Texas A&M University (TAMU) Tissue Chip Testing Center.

Figure 30: Representative example of the T_2 fitting in the cerebrospinal fluid (●), gray matter (○), and white matter (▼). The image in the top right corner shows the calculated T_2 map of the rat brain with the location of the fitted voxels (pointed to using the black lines). Each point in the graph represents the actual image intensity value (in institutional units, y-axis) in the given area of the echo image with the corresponding echo time (x-axis). A total of 16 echo images with 15 msec echo spacing were used. Black curves on the graph show the best fit function using the equation M 5 M0*exp(–TE/T_2), where M is the image intensity, M_0 is the image intensity of proton density image (at echo time = 0 msec), TE is echo time, and T_2 is the sought parameter. The T_2 map is color-coded according to the color scale on the right, each voxel in the map representing the observed quantitative T_2 relaxation in milliseconds.

FIGURE 31: Representative T_2 maps for animals which showed signs of brain alterations in response to treatment with known neurotoxicants. Control animal treated with saline (2 ml/kg, once) 48 h before imaging; 3-nitropropionic acid was administered in a dose of 20 mg/kg, s.c., daily for 3 days, imaging on day 4; hexachlorophene was administered in a dose of 30 mg/kg, p.o. daily for 5 days, imaging on day 6; kainic acid was administered in a dose of 10 mg/kg, i.p., once, imaging on day 3; domoic acid was administered in a dose of 2 mg/kg, i.p., once, imaging on day 3; pyrithiamine was administered in a dose of 0.25 mg/kg i.p. daily for 2 weeks + thiamine free diet, imaging at 4 weeks; trimethyltin was administered in a dose of 12 mg/kg, i.p., once, imaging at 3 weeks.

Figure 32: T_2 maps of a representative animal obtained at various time points (days, shown above each map) after hexachlorophene treatment. A single slice out of a 24-slice pack is shown. The color scale (on the right) relates to the actual T_2 relaxation time (in milliseconds) for each voxel. Note the maximum T_2 values in the white matter occurred at day 6. T_2 maps of the control subjects (not shown) were not different from T_2 maps at time 0.

Figure 33: Performance of iBiopsy® in advanced fibrosis prediction (*CRN=Clinical Research Network; ROC=receiver operating characteristic*).

Figure 34: Impact of fibrosis burden on HCC recurrence.

Figure 35: Modulation of Exposure-Response Relationships: Mechanisms.

Figure 36: A Type of Study Design Used to Assess Microbiome Influences on Toxicity (Many Variations on This Design Have Been Used).

Figure 37: A tiered approach for the risk assessment including the gut microbiome.

RReferences

1.   Allan J, Belz S, Hoeveler A, Hugas M, Okuda H, Patri A, Rauscher H, Silva P, Slikker W, Sokull-Kluettgen B, Tong W, Anklam E. Regulatory landscape of nanotechnology and nanoplastics from a global perspective. *Regul Toxicol Pharmacol* 2021;**122**:104885

2.   Honma M. An assessment of mutagenicity of chemical substances by (quantitative) structure-activity relationship. *Genes Environ* 2020;**42**:23

3.   Yamamoto E, Taquahashi Y, Kuwagata M, Saito H, Matsushita K, Toyoda T, Sato F, Kitajima S, Ogawa K, Izutsu K-i, Saito Y, Hirabayashi Y, Iimura Y, Honma M, Okuda H, Goda Y. Visualizing the spatial localization of ciclesonide and its metabolites in rat lungs after inhalation of 1-μm aerosol of ciclesonide by desorption electrospray ionization-time of flight mass spectrometry imaging. *International Journal of Pharmaceutics* 2021;**595**:120241

4.   Lambert D, Pightling A, Griffiths E, Van Domselaar G, Evans P, Berthelet S, Craig D, Chandry PS, Stones R, Brinkman F, Angers-Loustau A, Kreysa J, Tong W, Blais B. Baseline practices for the application of genomic data supporting regulatory food safety. *J AOAC Int* 2017;**100**:721-31

5.   Blais BW, Tapp K, Dixon M, Carrillo CD. Genomically informed strain-specific recovery of shiga toxin-producing escherichia coli during foodborne illness outbreak investigations. *J Food Prot* 2019;**82**:39-44

6.   Rott ME, Kesanakurti P, Berwarth C, Rast H, Boyes I, Phelan J, Jelkmann W. Discovery of negative-sense RNA viruses in trees infected with apple rubbery wood disease by next-generation sequencing. *Plant Dis* 2018;**102**:1254-63

7.   Lung O, Fisher M, Erickson A, Nfon C, Ambagala A. Fully automated and integrated multiplex detection of high consequence livestock viral genomes on a microfluidic platform.

Transbound Emerg Dis 2019;**66**:144-55

8. Authority EFS. Modern methodologies and tools for human hazard assessment of chemicals. *EFSA Journal* 2014;**12**:3638

9. Ball R, Robb M, Anderson SA, Dal Pan G. The FDA's sentinel initiative--A comprehensive approach to medical product surveillance. *Clin Pharmacol Ther* 2016;**99**:265-8

10. Platt R, Brown JS, Robb M, McClellan M, Ball R, Nguyen MD, Sherman RE. The FDA sentinel initiative - An evolving national resource. *N Engl J Med* 2018;**379**:2091-93

11. Brown JS, Maro JC, Nguyen M, Ball R. Using and improving distributed data networks to generate actionable evidence: the case of real-world outcomes in the Food and Drug Administration's sentinel system. *J Am Med Inform Assoc* 2020;**27**:793-97

12. Ball R, Toh S, Nolan J, Haynes K, Forshee R, Botsis T. Evaluating automated approaches to anaphylaxis case classification using unstructured data from the FDA sentinel system. *Pharmacoepidemiol Drug Saf* 2018;**27**:1077-84

13. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, Shafi S, Johnson DH, Mitter R, Rosenthal R, Salm M, Horswell S, Escudero M, Matthews N, Rowan A, Chambers T, Moore DA, Turajlic S, Xu H, Lee SM, Forster MD, Ahmad T, Hiley CT, Abbosh C, Falzon M, Borg E, Marafioti T, Lawrence D, Hayward M, Kolvekar S, Panagiotopoulos N, Janes SM, Thakrar R, Ahmed A, Blackhall F, Summers Y, Shah R, Joseph L, Quinn AM, Crosbie PA, Naidu B, Middleton G, Langman G, Trotter S, Nicolson M, Remmen H, Kerr K, Chetty M, Gomersall L, Fennell DA, Nakas A, Rathinam S, Anand G, Khan S, Russell P, Ezhil V, Ismail B, Irvin-Sellers M, Prakash V, Lester JF, Kornaszewska M, Attanoos R, Adams H, Davies H, Dentro S, Taniere P, O'Sullivan B, Lowe HL, Hartley JA, Iles N, Bell H, Ngai Y, Shaw JA, Herrero J, Szallasi Z, Schwarz RF, Stewart A, Quezada SA, Le Quesne J, Van Loo P, Dive C, Hackshaw A, Swanton C. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med* 2017;**376**:2109-21

14. AbdulJabbar K, Raza SEA, Rosenthal R, Jamal-Hanjani M, Veeriah S, Akarca A, Lund T, Moore DA, Salgado R, Al Bakir M, Zapata L, Hiley CT, Officer L, Sereno M, Smith CR, Loi S, Hackshaw A, Marafioti T, Quezada SA, McGranahan N, Le Quesne J, Swanton C, Yuan Y. Geospatial immune variability illuminates differential evolution of lung adenocarcinoma. *Nat Med* 2020;**26**:1054-62

15. Pennycuick A, Teixeira VH, AbdulJabbar K, Raza SEA, Lund T, Akarca AU, Rosenthal R, Kalinke L, Chandrasekharan DP, Pipinikas CP, Lee-Six H, Hynds RE, Gowers KHC, Henry JY, Millar FR, Hagos YB, Denais C, Falzon M, Moore DA, Antoniou S, Durrenberger PF, Furness AJ, Carroll B, Marceaux C, Asselin-Labat ML, Larson W, Betts C, Coussens LM, Thakrar RM, George J, Swanton C, Thirlwell C, Campbell PJ, Marafioti T, Yuan Y, Quezada SA, McGranahan N, Janes SM. Immune surveillance in clinical regression of preinvasive squamous cell lung cancer. *Cancer*

Discov 2020;**10**:1489-99

16.   Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc* 2020;**27**:491-97

17.   Viceconti M, Pappalardo F, Rodriguez B, Horner M, Bischoff J, Musuamba Tshinanu F. In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products. *Methods* 2021;**185**:120-27

18.   Consortium SM-I. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the sequencing quality control consortium. *Nat Biotechnol* 2014;**32**:903-14

19.   Tong L, Wu PY, Phan JH, Hassazadeh HR, Consortium S, Tong W, Wang MD. Impact of RNA-seq data analysis algorithms on gene expression estimation and downstream prediction. *Sci Rep* 2020;**10**:17925

20.   Tong L, Mitchel J, Chatlin K, Wang MD. Deep learning based feature-level integration of multi-omics data for breast cancer patients survival analysis. *BMC Med Inform Decis Mak* 2020;**20**:225

21.   Tong L, Wu H, Wang MD. Integrating multi-omics data by learning modality invariant representations for improved prediction of overall survival of cancer. *Methods* 2020;

22.   Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrera F. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 2020;**58**:82-115

23.   Boulemtafes A, Derhab A, Challal Y. A review of privacy-preserving techniques for deep learning. *Neurocomputing* 2020;**384**:21-45

24.   Hardwick SA, Deveson IW, Mercer TR. Reference standards for next-generation sequencing. *Nat Rev Genet* 2017;**18**:473-84

25.   Zook JM, Catoe D, McDaniel J, Vang L, Spies N, Sidow A, Weng Z, Liu Y, Mason CE, Alexander N, Henaff E, McIntyre AB, Chandramohan D, Chen F, Jaeger E, Moshrefi A, Pham K, Stedman W, Liang T, Saghbini M, Dzakula Z, Hastie A, Cao H, Deikus G, Schadt E, Sebra R, Bashir A, Truty RM, Chang CC, Gulbahce N, Zhao K, Ghosh S, Hyland F, Fu Y, Chaisson M, Xiao C, Trow J, Sherry ST, Zaranek AW, Ball M, Bobe J, Estep P, Church GM, Marks P, Kyriazopoulou-Panagiotopoulou S, Zheng GX, Schnall-Levin M, Ordonez HS, Mudivarti PA, Giorda K, Sheng Y, Rypdal KB, Salit M. Extensive sequencing of seven human genomes to characterize benchmark reference materials. *Sci Data* 2016;**3**:160025

26.   Deveson IW, Chen WY, Wong T, Hardwick SA, Andersen SB, Nielsen LK, Mattick JS, Mercer TR. Representing genetic variation with synthetic DNA standards. *Nat Methods* 2016;**13**:784-91

27.   Cescon D, Bratman S, Chan S, Siu L. Circulating tumor DNA and liquid biopsy in oncology. *Nature Cancer* 2020;**1**:276-90

28.   Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, Farrar K, Park BK, Breckenridge AM. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;**329**:15-19

29.   Davies EC, Green CF, Taylor S, Williamson PR, Mottram DR, Pirmohamed M. Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. *PLoS One* 2009;**4**:e4439

30.   Breiteneder H, Peng YQ, Agache I, Diamant Z, Eiwegger T, Fokkens WJ, Traidl-Hoffmann C, Nadeau K, O'Hehir RE, O'Mahony L, Pfaar O, Torres MJ, Wang DY, Zhang L, Akdis CA. Biomarkers for diagnosis and prediction of therapy responses in allergic diseases and asthma. *Allergy* 2020;**75**:3039-68

31.   Evans WE, Relling MV. Moving towards individualized medicine with pharmacogenomics. *Nature* 2004;**429**:464-8

32.   Turner RM, Newman WG, Bramon E, McNamee CJ, Wong WL, Misbah S, Hill S, Caulfield M, Pirmohamed M. Pharmacogenomics in the UK National Health Service: opportunities and challenges. *Pharmacogenomics* 2020;**21**:1237-46

33.   van der Wouden CH, Cambon-Thomsen A, Cecchin E, Cheung KC, Dávila-Fajardo CL, Deneer VH, Dolžan V, Ingelman-Sundberg M, Jönsson S, Karlsson MO, Kriek M, Mitropoulou C, Patrinos GP, Pirmohamed M, Samwald M, Schaeffeler E, Schwab M, Steinberger D, Stingl J, Sunder-Plassmann G, Toffoli G, Turner RM, van Rhenen MH, Swen JJ, Guchelaar HJ. Implementing pharmacogenomics in Europe: Design and implementation strategy of the ubiquitous pharmacogenomics consortium. *Clin Pharmacol Ther* 2017;**101**:341-58

34.   Pirmohamed M, Burnside G, Eriksson N, Jorgensen AL, Toh CH, Nicholson T, Kesteven P, Christersson C, Wahlström B, Stafberg C, Zhang JE, Leathart JB, Kohnke H, Maitland-van der Zee AH, Williamson PR, Daly AK, Avery P, Kamali F, Wadelius M. A randomized trial of genotype-guided dosing of warfarin. *N Engl J Med* 2013;**369**:2294-303

35.   Jorgensen AL, Prince C, Fitzgerald G, Hanson A, Downing J, Reynolds J, Zhang JE, Alfirevic A, Pirmohamed M. Implementation of genotype-guided dosing of warfarin with point-of-care genetic testing in three UK clinics: a matched cohort study. *BMC Medicine* 2019;**17**:76

36.   Kimmel SE, French B, Kasner SE, Johnson JA, Anderson JL, Gage BF, Rosenberg YD, Eby CS, Madigan RA, McBane RB, Abdel-Rahman SZ, Stevens SM, Yale S, Mohler ER, Fang MC, Shah V, Horenstein RB, Limdi NA, Muldowney JAS, Gujral J, Delafontaine P, Desnick RJ, Ortel TL, Billett HH, Pendleton RC, Geller NL, Halperin JL, Goldhaber SZ, Caldwell MD, Califf RM, Ellenberg JH. A pharmacogenetic versus a clinical algorithm for warfarin dosing. *New England Journal of Medicine* 2013;**369**:2283-93

37.   Pirmohamed M, Ostrov DA, Park BK. New genetic findings lead the way to a better understanding of fundamental mechanisms of drug hypersensitivity. *J Allergy Clin Immunol* 2015;**136**:236-44

38.   Beger RD, Sun J, Schnackenberg LK. Metabolomics approaches for discovering biomarkers of drug-induced hepatotoxicity and nephrotoxicity. *Toxicol Appl Pharmacol* 2010;**243**:154-66

39.   Viant MR, Ebbels TMD, Beger RD, Ekman DR, Epps DJT, Kamp H, Leonards PEG, Loizou GD, MacRae JI, van Ravenzwaay B, Rocca-Serra P, Salek RM, Walk T, Weber RJM. Use cases, best practice and reporting standards for metabolomics in regulatory toxicology. *Nat Commun* 2019;**10**:3041

40.   Evans AM, O'Donovan C, Playdon M, Beecher C, Beger RD, Bowden JA, Broadhurst D, Clish CB, Dasari S, Dunn WB, Griffin JL, Hartung T, Hsu PC, Huan T, Jans J, Jones CM, Kachman M, Kleensang A, Lewis MR, Monge ME, Mosley JD, Taylor E, Tayyari F, Theodoridis G, Torta F, Ubhi BK, Vuckovic D. Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC-MS based untargeted metabolomics practitioners. *Metabolomics* 2020;**16**:113

41.   Cao Z, Kamlage B, Wagner-Golbs A, Maisha M, Sun J, Schnackenberg LK, Pence L, Schmitt TC, Daniels JR, Rogstad S, Beger RD, Yu LR. An integrated analysis of metabolites, peptides, and inflammation biomarkers for assessment of preanalytical variability of human plasma. *J Proteome Res* 2019;**18**:2411-21

42.   Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, Brennan L, Wishart DS, Oresic M, Hankemeier T, Broadhurst DI, Lane AN, Suhre K, Kastenmüller G, Sumner SJ, Thiele I, Fiehn O, Kaddurah-Daouk R. Metabolomics enables precision medicine: "A White Paper, Community Perspective". *Metabolomics* 2016;**12**:149

43.   Mussap M, Zaffanello M, Fanos V. Metabolomics: A challenge for detecting and monitoring inborn errors of metabolism. *Ann Transl Med* 2018;**6**:338

44.   Miller MJ, Kennedy AD, Eckhart AD, Burrage LC, Wulff JE, Miller LA, Milburn MV, Ryals JA, Beaudet AL, Sun Q, Sutton VR, Elsea SH. Untargeted metabolomic analysis for the clinical screening of inborn errors of metabolism. *J Inherit Metab Dis* 2015;**38**:1029-39

45.   Clayton TA, Lindon JC, Cloarec O, Antti H, Charuel C, Hanton G, Provost JP, Le Net JL, Baker D, Walley RJ, Everett JR, Nicholson JK. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature* 2006;**440**:1073-7

46.   Clayton TA, Baker D, Lindon JC, Everett JR, Nicholson JK. Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A* 2009;**106**:14728-33

47.   Beger RD, Bhattacharyya S, Yang X, Gill PS, Schnackenberg LK, Sun J, James LP.

Translational biomarkers of acetaminophen-induced acute liver injury. *Arch Toxicol* 2015;**89**:1497-522

48.   Li YY, Ghanbari R, Pathmasiri W, McRitchie S, Poustchi H, Shayanrad A, Roshandel G, Etemadi A, Pollock JD, Malekzadeh R, Sumner SCJ. Untargeted Metabolomics: Biochemical perturbations in golestan cohort study opium users inform intervention strategies. *Front Nutr* 2020;**7**:584585

49.   Ghanbari R, Li Y, Pathmasiri W, McRitchie S, Etemadi A, Pollock JD, Poustchi H, Rahimi-Movaghar A, Amin-Esmaeili M, Roshandel G, Shayanrad A, Abaei B, Malekzadeh R, Sumner SCJ. Metabolomics reveals biomarkers of opioid use disorder. *Transl Psychiatry* 2021;**11**:103

50.   Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res* 2018;**46**:D1074-d82

51.   Wishart D, Arndt D, Pon A, Sajed T, Guo AC, Djoumbou Y, Knox C, Wilson M, Liang Y, Grant J, Liu Y, Goldansaz SA, Rappaport SM. T3DB: The toxic exposome database. *Nucleic Acids Res* 2015;**43**:D928-34

52.   Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, Sayeeda Z, Lo E, Assempour N, Berjanskii M, Singhal S, Arndt D, Liang Y, Badran H, Grant J, Serra-Cayuela A, Liu Y, Mandal R, Neveu V, Pon A, Knox C, Wilson M, Manach C, Scalbert A. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res* 2018;**46**:D608-d17

53.   Wishart DS, Bartok B, Oler E, Liang KYH, Budinski Z, Berjanskii M, Guo A, Cao X, Wilson M. MarkerDB: An online database of molecular biomarkers. *Nucleic Acids Res* 2021;**49**:D1259-d67

54.   Neveu V, Nicolas G, Salek RM, Wishart DS, Scalbert A. Exposome-Explorer 2.0: An update incorporating candidate dietary biomarkers and dietary associations with cancer risk. *Nucleic Acids Res* 2020;**48**:D908-d12

55.   Djoumbou-Feunang Y, Fiamoncini J, Gil-de-la-Fuente A, Greiner R, Manach C, Wishart DS. BioTransformer: A comprehensive computational tool for small molecule metabolism prediction and metabolite identification. *J Cheminform* 2019;**11**:2

56.   Salek RM, Neumann S, Schober D, Hummel J, Billiau K, Kopka J, Correa E, Reijmers T, Rosato A, Tenori L, Turano P, Marin S, Deborde C, Jacob D, Rolin D, Dartigues B, Conesa P, Haug K, Rocca-Serra P, O'Hagan S, Hao J, Vliet Mv, Sysi-Aho M, Ludwig C, Bouwman J, Cascante M, Ebbels T, Griffin JL, Moing A, Nikolski M, Oresic M, Sansone S-A, Viant MR, Goodacre R, GÃ¼nther UL, Hankemeier T, Luchinat C, Walther D, Steinbeck C. Erratum to: COordination of Standards in MetabOlomicS (COSMOS): Facilitating integrated metabolomics data access.

Metabolomics 2015;**11**:1598--99

57.   Salek RM, Haug K, Steinbeck C. Dissemination of metabolomics results: role of MetaboLights and COSMOS. *GigaScience* 2013;**2**:8

58.   Rijswijk Mv, Beirnaert C, Caron C, Cascante M, Dominguez V, Dunn WB, Ebbels TMD, Giacomoni F, Gonzalez-Beltran A, Hankemeier T, Haug K, Izquierdo-Garcia JL, Jimenez RC, Jourdan F, Kale N, Klapa MI, Kohlbacher O, Koort K, Kultima K, Corguillé GL, Moreno P, Moschonas NK, Neumann S, O'Donovan C, Reczko M, Rocca-Serra P, Rosato A, Salek RM, Sansone S-A, Satagopam V, Schober D, Shimmo R, Spicer RA, Spjuth O, Thévenot EA, Viant MR, Weber RJM, Willighagen EL, Zanetti G, Steinbeck C. The future of metabolomics in ELIXIR [version 2; referees: 3 approved]. 2017;

59.   Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, Overvad K, Tjonneland A, Clavel-Chapelon F, Thiebaut A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-De-Mesquita HB, Peeters PH, Lund E, Engeset D, Gonzalez CA, Barricarte A, Berglund G, Hallmans G, Day NE, Key TJ, Kaaks R, Saracci R. European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002;**5**:1113-24

60.   Hoffmann N, Hartler J, Ahrends R. jmzTab-M: A reference parser, writer, and validator for the proteomics standards initiative mzTab 2.0 metabolomics standard. *Anal Chem* 2019;**91**:12615-18

61.   Morello A, Sadelain M, Adusumilli PS. Mesothelin-Targeted CARs: Driving T cells to solid tumors. *Cancer Discov* 2016;**6**:133-46

62.   Neelapu SS, Locke FL, Bartlett NL, Lekakis LJ, Miklos DB, Jacobson CA, Braunschweig I, Oluwole OO, Siddiqi T, Lin Y, Timmerman JM, Stiff PJ, Friedberg JW, Flinn IW, Goy A, Hill BT, Smith MR, Deol A, Farooq U, McSweeney P, Munoz J, Avivi I, Castro JE, Westin JR, Chavez JC, Ghobadi A, Komanduri KV, Levy R, Jacobsen ED, Witzig TE, Reagan P, Bot A, Rossi J, Navale L, Jiang Y, Aycock J, Elias M, Chang D, Wiezorek J, Go WY. Axicabtagene ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. *N Engl J Med* 2017;**377**:2531-44

63.   Ruella M, Kenderian SS. Next-Generation chimeric antigen receptor T-cell therapy: going off the shelf. *BioDrugs* 2017;**31**:473-81

64.   Hashemi S, Fransen MF, Niemeijer A, Ben Taleb N, Houda I, Veltman J, Becker-Commissaris A, Daniels H, Crombag L, Radonic T, Jongeneel G, Tarasevych S, Looysen E, van Laren M, Tiemessen M, van Diepen V, Maassen-van den Brink K, Thunnissen E, Bahce I. Surprising impact of stromal TIL's on immunotherapy efficacy in a real-world lung cancer study. *Lung Cancer* 2021;**153**:81-89

65.   Grigor EJM, Fergusson D, Kekre N, Montroy J, Atkins H, Seftel MD, Daugaard M, Presseau J, Thavorn K, Hutton B, Holt RA, Lalu MM. Risks and benefits of chimeric antigen

receptor T-cell (CAR-T) therapy in cancer: A systematic review and meta-analysis. *Transfus Med Rev* 2019;**33**:98-110

66. Wang X, Qi Z, Wei H, Tian Z, Sun R. Characterization of human B cells in umbilical cord blood-transplanted NOD/SCID mice. *Transpl Immunol* 2012;**26**:156-62

67. Hu Y, Wu Z, Luo Y, Shi J, Yu J, Pu C, Liang Z, Wei G, Cui Q, Sun J, Jiang J, Xie J, Tan Y, Ni W, Tu J, Wang J, Jin A, Zhang H, Cai Z, Xiao L, Huang H. Potent anti-leukemia activities of chimeric antigen receptor-modified T cells against CD19 in chinese patients with relapsed/refractory acute lymphocytic leukemia. *Clin Cancer Res* 2017;**23**:3297-306

68. Wen H, Qu Z, Yan Y, Pu C, Wang C, Jiang H, Hou T, Huo Y. Preclinical safety evaluation of chimeric antigen receptor-modified T cells against CD19 in NSG mice. *Ann Transl Med* 2019;**7**:735

69. Kimura H, Sakai Y, Fujii T. Organ/body-on-a-chip based on microfluidic technology for drug discovery. *Drug Metab Pharmacokinet* 2018;**33**:43-48

70. Tetsuka K, Ohbuchi M, Kawabe T, Goto T, Kiyonaga F, Takama K, Yamazaki S, Fujimori A. Reconstituted human organ models as a translational tool for human organ response: Definition, expectations, cases, and strategies for implementation in drug discovery and development. *Biol Pharm Bull* 2020;**43**:375-83

71. Ishida S. Organs-on-a-chip: Current applications and consideration points for in vitro ADME-Tox studies. *Drug Metab Pharmacokinet* 2018;**33**:49-54

72. Satoh T, Sugiura S, Shin K, Onuki-Nagasaki R, Ishida S, Kikuchi K, Kakiki M, Kanamori T. A multi-throughput multi-organ-on-a-chip system on a plate formatted pneumatic pressure-driven medium circulation platform. *Lab Chip* 2017;**18**:115-25

73. Sano E, Mori C, Matsuoka N, Ozaki Y, Yagi K, Wada A, Tashima K, Yamasaki S, Tanabe K, Yano K, Torisawa YS. Tetrafluoroethylene-Propylene elastomer for fabrication of microfluidic organs-on-chips resistant to drug absorption. *Micromachines (Basel)* 2019;**10**

74. Ishida S. Requirements for designing organ-on-a-chip platforms to model the pathogenesis of liver disease. In: Hoeng J, Bovard D, Peitsch M, editors. Organ-on-a-chip academic press, 2019. pp. 181-213

75. Ribeiro AJS, Yang X, Patel V, Madabushi R, Strauss DG. Liver Microphysiological systems for predicting and evaluating drug effects. *Clin Pharmacol Ther* 2019;**106**:139-47

76. Rouse R, Kruhlak N, Weaver J, Burkhart K, Patel V, Strauss DG. Translating new science into the srug review process: The US FDA's division of applied regulatory science. *Ther Innov Regul Sci* 2018;**52**:244-55

77. Dame K, Ribeiro AJ. Microengineered systems with iPSC-derived cardiac and hepatic cells to evaluate drug adverse effects. *Exp Biol Med (Maywood)* 2020:1535370220959598

78. Ribeiro AJS, Guth BD, Engwall M, Eldridge S, Foley CM, Guo L, Gintant G, Koerner J,

Parish ST, Pierson JB, Brock M, Chaudhary KW, Kanda Y, Berridge B. Considerations for an in vitro, cell-based testing platform for detection of drug-induced inotropic effects in early drug development. Part 2: designing and fabricating microsystems for assaying cardiac contractility with physiological relevance using human iPSC-cardiomyocytes. *Front Pharmacol* 2019;**10**:934

79.   Rubiano A, Indapurkar A, Yokosawa R, Miedzik A, Rosenzweig B, Arefin A, Moulin CM, Dame K, Hartman N, Volpe DA, Matta MK, Hughes DJ, Strauss DG, Kostrzewski T, Ribeiro AJS. Characterizing the reproducibility in using a liver microphysiological system for assaying arug toxicity, metabolism and accumulation. *Clin Transl Sci* 2020;

80.   Roberts RA, Ganey PE, Ju C, Kamendulis LM, Rusyn I, Klaunig JE. Role of the kupffer cell in mediating hepatic toxicity and carcinogenesis. *Toxicol Sci* 2007;**96**:2-15

81.   Miller JM, Meki MH, Ou Q, George SA, Gams A, Abouleisa RRE, Tang XL, Ahern BM, Giridharan GA, El-Baz A, Hill BG, Satin J, Conklin DJ, Moslehi J, Bolli R, Ribeiro AJS, Efimov IR, Mohamed TMA. Heart slice culture system reliably demonstrates clinical drug-related cardiotoxicity. *Toxicol Appl Pharmacol* 2020;**406**:115213

82.   Loskill P, Marcus SG, Mathur A, Reese WM, Healy KE. μOrgano: A Lego®-Like plug & play system for modular multi-organ-chips. *PLoS One* 2015;**10**:e0139587

83.   Rampe D, Wible B, Brown AM, Dage RC. Effects of terfenadine and its metabolites on a delayed rectifier K+ channel cloned from human heart. *Mol Pharmacol* 1993;**44**:1240-5

84.   Mannhardt I, Breckwoldt K, Letuffe-Brenière D, Schaaf S, Schulz H, Neuber C, Benzin A, Werner T, Eder A, Schulze T, Klampe B, Christ T, Hirt MN, Huebner N, Moretti A, Eschenhagen T, Hansen A. Human engineered heart tissue: analysis of contractile force. *Stem Cell Reports* 2016;**7**:29-42

85.   Mastrangeli M, Millet S, Mummery C, Loskill P, Braeken D, Eberle W, Cipriano M, Fernandez L, Graef M, Gidrol X, Picollet-D'Hahan N, Van Meer B, Ochoa I, Schutte M, Van den Eijnden-van Raaij J. Building blocks for a european organ-on-chip roadmap. *Altex* 2019;**36**:481-92

86.   Mastrangeli M, Millet S, Orchid Partners T, Van den Eijnden-van Raaij J. Organ-on-chip in development: Towards a roadmap for organs-on-chip. *Altex* 2019;**36**:650-68

87.   Powers MJ, Domansky K, Kaazempur-Mofrad MR, Kalezi A, Capitano A, Upadhyaya A, Kurzawski P, Wack KE, Stolz DB, Kamm R, Griffith LG. A microfabricated array bioreactor for perfused 3D liver culture. *Biotechnol Bioeng* 2002;**78**:257-69

88.   Sin A, Chin KC, Jamil MF, Kostov Y, Rao G, Shuler ML. The design and fabrication of three-chamber microscale cell culture analog devices with integrated dissolved oxygen sensors. *Biotechnol Prog* 2004;**20**:338-45

89.   Marx U. How drug development of the 21st century could benefit from human micro-organoid in-vitro technologies. In: Marx U, S. V, editors. Drug testing in vitro: Breakthroughs and

trends in cell culture technology: Wiley, 2006. pp. 181-213

90.   Marx U, Walles H, Hoffmann S, Lindner G, Horland R, Sonntag F, Klotzbach U, Sakharov D, Tonevitsky A, Lauster R. 'Human-on-a-chip' developments: a translational cutting-edge alternative to systemic safety assessment and efficiency evaluation of substances in laboratory animals and man? *Altern Lab Anim* 2012;**40**:235-57

91.   Huh D, Matthews BD, Mammoto A, Montoya-Zavala M, Hsin HY, Ingber DE. Reconstituting organ-level lung functions on a chip. *Science* 2010;**328**:1662-8

92.   Livingston CA, Fabre KM, Tagle DA. Facilitating the commercialization and use of organ platforms generated by the microphysiological systems (Tissue Chip) program through public-private partnerships. *Comput Struct Biotechnol J* 2016;**14**:207-10

93.   Park SE, Georgescu A, Huh D. Organoids-on-a-chip. *Science* 2019;**364**:960-65

94.   Marx U, Akabane T, Andersson TB, Baker E, Beilmann M, Beken S, Brendler-Schwaab S, Cirit M, David R, Dehne EM, Durieux I, Ewart L, Fitzpatrick SC, Frey O, Fuchs F, Griffith LG, Hamilton GA, Hartung T, Hoeng J, Hogberg H, Hughes DJ, Ingber DE, Iskandar A, Kanamori T, Kojima H, Kuehnl J, Leist M, Li B, Loskill P, Mendrick DL, Neumann T, Pallocca G, Rusyn I, Smirnova L, Steger-Hartmann T, Tagle DA, Tonevitsky A, Tsyb S, Trapecar M, Van de Water B, Van den Eijnden-van Raaij J, Vulto P, Watanabe K, Wolf A, Zhou X, Roth A. Biology-inspired microphysiological systems to advance patient benefit and animal welfare in drug development. *Altex* 2020;**37**:365-94

95.   Begley CG. Six red flags for suspect work. *Nature* 2013;**497**:433-4

96.   Lin N, Zhou X, Geng X, Drewell C, Hübner J, Li Z, Zhang Y, Xue M, Marx U, Li B. Repeated dose multi-drug testing using a microfluidic chip-based coculture of human liver and kidney proximal tubules equivalents. *Sci Rep* 2020;**10**:8879

97.   Dehne E-M, Marx U. The universal physiological template—a system to advance medicines. *Current Opinion in Toxicology* 2020;**23-24**:1-5

98.   Beilmann M, Boonen H, Czich A, Dear G, Hewitt P, Mow T, Newham P, Oinonen T, Pognan F, Roth A, Valentin JP, Van Goethem F, Weaver RJ, Birk B, Boyer S, Caloni F, Chen AE, Corvi R, Cronin MTD, Daneshian M, Ewart LC, Fitzgerald RE, Hamilton GA, Hartung T, Kangas JD, Kramer NI, Leist M, Marx U, Polak S, Rovida C, Testai E, Van der Water B, Vulto P, Steger-Hartmann T. Optimizing drug discovery by investigative toxicology: Current and future trends. *Altex* 2019;**36**:289-313

99.   Pamies D, Bal-Price A, Chesné C, Coecke S, Dinnyes A, Eskes C, Grillari R, Gstraunthaler G, Hartung T, Jennings P, Leist M, Martin U, Passier R, Schwamborn JC, Stacey GN, Ellinger-Ziegelbauer H, Daneshian M. Advanced good cell culture practice for human primary, stem cell-derived and organoid models as well as microphysiological systems. *Altex* 2018;**35**:353-78

100.   Sieber S, Wirth L, Cavak N, Koenigsmark M, Marx U, Lauster R, Rosowski M. Bone

marrow-on-a-chip: Long-term culture of human haematopoietic stem cells in a three-dimensional microfluidic environment. *J Tissue Eng Regen Med* 2018;**12**:479-89

101. Schoon J, Hesse B, Rakow A, Ort MJ, Lagrange A, Jacobi D, Winter A, Huesker K, Reinke S, Cotte M, Tucoulou R, Marx U, Perka C, Duda GN, Geissler S. Metal-Specific biomaterial accumulation in human peri-implant bone and bone Marrow. *Adv Sci (Weinh)* 2020;**7**:2000412

102. Lam J, Bellayr IH, Marklein RA, Bauer SR, Puri RK, Sung KE. Functional profiling of chondrogenically induced multipotent stromal cell aggregates reveals transcriptomic and emergent morphological phenotypes predictive of differentiation capacity. *Stem Cells Transl Med* 2018;**7**:664-75

103. Singh P, Schwarzbauer JE. Fibronectin and stem cell differentiation - Lessons from chondrogenesis. *J Cell Sci* 2012;**125**:3703-12

104. Kim S, Lee H, Chung M, Jeon NL. Engineering of functional, perfusable 3D microvascular networks on a chip. *Lab Chip* 2013;**13**:1489-500

105. Barrett, R., Ornelas, L., Yeager, N., et al. 2014. Reliable generation of induced pluripotent stem cells from human lymphoblastoid cell lines. *Stem cells translational medicine* 3(12), pp. 1429–1434

106. Laperle, A.H., Sances, S., Yucer, N., et al. 2020. iPSC modeling of young-onset Parkinson's disease reveals a molecular signature of disease and novel therapeutic candidates.

107. *Nature Medicine* 26(2), pp. 289–299

Workman, M.J. and Svendsen, C.N. 2020. Recent advances in human iPSC-derived models of the blood-brain barrier. *Fluids and barriers of the CNS* 17(1), p. 30

108. Vatine GD, Barrile R, Workman MJ, Sances S, Barriga BK, Rahnama M, Barthakur S, Kasendra M, Lucchesi C, Kerns J, Wen N, Spivia WR, Chen Z, Van Eyk J, Svendsen CN. Human iPSC-derived blood-brain barrier chips enable disease modeling and personalized medicine applications. *Cell Stem Cell* 2019;**24**:995-1005.e6

109. Lim RG, Quan C, Reyes-Ortiz AM, Lutz SE, Kedaigle AJ, Gipson TA, Wu J, Vatine GD, Stocksdale J, Casale MS, Svendsen CN, Fraenkel E, Housman DE, Agalliu D, Thompson LM. Huntington's disease iPSC-derived brain microvascular endothelial cells reveal WNT-mediated angiogenic and blood-brain barrier deficits. *Cell Rep* 2017;**19**:1365-77

110. Sances S, Ho R, Vatine G, West D, Laperle A, Meyer A, Godoy M, Kay PS, Mandefro B, Hatata S, Hinojosa C, Wen N, Sareen D, Hamilton GA, Svendsen CN. Human iPSC-derived endothelial cells and microengineered organ-chip enhance neuronal development. *Stem Cell Reports* 2018;**10**:1222-36

111. Toombs J, Panther L, Ornelas L, Liu C, Gomez E, Martín-Ibáñez R, Cox SR, Ritchie SJ, Harris SE, Taylor A, Redmond P, Russ TC, Murphy L, Cooper JD, Burr K, Selvaraj BT, Browne

C, Svendsen CN, Cowley SA, Deary IJ, Chandran S, Spires-Jones TL, Sareen D. Generation of twenty four induced pluripotent stem cell lines from twenty four members of the Lothian Birth Cohort 1936. *Stem Cell Res* 2020;**46**:101851

112.   Sharma A, Garcia G, Jr., Wang Y, Plummer JT, Morizono K, Arumugaswami V, Svendsen CN. Human iPSC-derived cardiomyocytes are susceptible to SARS-CoV-2 infection. *Cell Rep Med* 2020;**1**:100052

113.   Sharma A, Sances S, Workman MJ, Svendsen CN. Multi-lineage human iPSC-derived platforms for disease modeling and drug discovery. *Cell Stem Cell* 2020;**26**:309-29

114.   Low LA, Sutherland M, Lumelsky N, Selimovic S, Lundberg MS, Tagle DA. Organs-on-a-Chip. *Adv Exp Med Biol* 2020;**1230**:27-42

115.   Fabre K, Berridge B, Proctor WR, Ralston S, Will Y, Baran SW, Yoder G, Van Vleet TR. Introduction to a manuscript series on the characterization and use of microphysiological systems (MPS) in pharmaceutical safety and ADME applications. *Lab Chip* 2020;**20**:1049-57

116.   Low LA, Tagle DA. Organs-on-chips: Progress, challenges, and future directions. *Exp Biol Med (Maywood)* 2017;**242**:1573-78

117.   Ewart L, Fabre K, Chakilam A, Dragan Y, Duignan DB, Eswaraka J, Gan J, Guzzie-Peck P, Otieno M, Jeong CG, Keller DA, de Morais SM, Phillips JA, Proctor W, Sura R, Van Vleet T, Watson D, Will Y, Tagle D, Berridge B. Navigating tissue chips from development to dissemination: A pharmaceutical industry perspective. *Exp Biol Med (Maywood)* 2017;**242**:1579-85

118.   Weber EJ, Chapron A, Chapron BD, Voellinger JL, Lidberg KA, Yeung CK, Wang Z, Yamaura Y, Hailey DW, Neumann T, Shen DD, Thummel KE, Muczynski KA, Himmelfarb J, Kelly EJ. Development of a microphysiological model of human kidney proximal tubule function. *Kidney Int* 2016;**90**:627-37

119.   Weber EJ, Lidberg KA, Wang L, Bammler TK, MacDonald JW, Li MJ, Redhair M, Atkins WM, Tran C, Hines KM, Herron J, Xu L, Monteiro MB, Ramm S, Vaidya V, Vaara M, Vaara T, Himmelfarb J, Kelly EJ. Human kidney on a chip assessment of polymyxin antibiotic nephrotoxicity. *JCI Insight* 2018;**3**

120.   Sakolish C, Weber EJ, Kelly EJ, Himmelfarb J, Mouneimne R, Grimm FA, House JS, Wade T, Han A, Chiu WA, Rusyn I. Technology transfer of the microphysiological systems: A case study of the human proximal tubule tissue chip. *Sci Rep* 2018;**8**:14882

121.   Vernetti LA, Senutovitch N, Boltz R, DeBiasio R, Shun TY, Gough A, Taylor DL. A human liver microphysiology platform for investigating physiology, drug safety, and disease models. *Exp Biol Med (Maywood)* 2016;**241**:101-14

122.   Sakolish C, Reese CE, Luo YS, Valdiviezo A, Schurdak ME, Gough A, Taylor DL, Chiu WA, Vernetti LA, Rusyn I. Analysis of reproducibility and robustness of a human microfluidic

four-cell liver acinus microphysiology system (LAMPS). *Toxicology* 2021;**448**:152651

123.   Moek KL, Giesen D, Kok IC, de Groot DJA, Jalving M, Fehrmann RSN, Lub-de Hooge MN, Brouwers AH, de Vries EGE. Theranostics using antibodies and antibody-related therapeutics. *J Nucl Med* 2017;**58**:83s-90s

124.   Barkhof F, Daams M, Scheltens P, Brashear HR, Arrighi HM, Bechten A, Morris K, McGovern M, Wattjes MP. An MRI rating scale for amyloid-related imaging abnormalities with edema or effusion. *AJNR Am J Neuroradiol* 2013;**34**:1550-5

125.   Pasquier F, Sadowsky C, Holstein A, Leterme Gle P, Peng Y, Jackson N, Fox NC, Ketter N, Liu E, Ryan JM. Two phase 2 multiple ascending-dose studies of vanutide cridificar (ACC-001) and QS-21 adjuvant in mild-to-moderate alzheimer's disease. *J Alzheimers Dis* 2016;**51**:1131-43

126.   Delnomdedieu M, Duvvuri S, Li DJ, Atassi N, Lu M, Brashear HR, Liu E, Ness S, Kupiec JW. First-In-Human safety and long-term exposure data for AAB-003 (PF-05236812) and biomarkers after intravenous infusions of escalating doses in patients with mild to moderate Alzheimer's disease. *Alzheimers Res Ther* 2016;**8**:12

127.   Salloway S, Sperling R, Gilman S, Fox NC, Blennow K, Raskind M, Sabbagh M, Honig LS, Doody R, van Dyck CH, Mulnard R, Barakos J, Gregg KM, Liu E, Lieberburg I, Schenk D, Black R, Grundman M. A phase 2 multiple ascending dose trial of bapineuzumab in mild to moderate Alzheimer disease. *Neurology* 2009;**73**:2061-70

128.   Sperling RA, Jack CR, Jr., Black SE, Frosch MP, Greenberg SM, Hyman BT, Scheltens P, Carrillo MC, Thies W, Bednar MM, Black RS, Brashear HR, Grundman M, Siemers ER, Feldman HH, Schindler RJ. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the alzheimer's association research roundtable workgroup. *Alzheimers Dement* 2011;**7**:367-85

129.   Gregoire SM, Chaudhary UJ, Brown MM, Yousry TA, Kallis C, Jäger HR, Werring DJ. The microbleed anatomical rating scale (MARS): Reliability of a tool to map brain microbleeds. *Neurology* 2009;**73**:1759-66

130.   Kenna JG, Waterton JC, Baudy A, Galetin A, Hines CDG, Hockings P, Patel M, Scotcher D, Sourbron S, Ziemian S, Schuetz G. Noninvasive preclinical and clinical imaging of liver transporter function relevant to drug-induced liver injury. In: Chen M, Will Y, editors. Drug-induced liver toxicity. New York, NY: Springer New York, 2018. pp. 627-51

131.   Bonnaventure P, Cusin F, Pastor CM. Hepatocyte concentrations of imaging compounds associated with transporter inhibition: Evidence in perfused rat livers. *Drug Metab Dispos* 2019;**47**:412-18

132.   Ulloa J, Stahl S, Woodhouse N, Halliday J, Parmar A, Holmes A, Barjat H, Hockings P. Effects of a single intravenous dose of Estradiol-17^2 D-glucuronide on biliary excretion:

Assessment with gadoxetate DCEMRI. Proc 18th Sci Meet Int Soc Magn Reson2010.

133. Ulloa JL, Stahl S, Yates J, Woodhouse N, Kenna JG, Jones HB, Waterton JC, Hockings PD. Assessment of gadoxetate DCE-MRI as a biomarker of hepatobiliary transporter inhibition. *NMR Biomed* 2013;**26**:1258-70

134. Karageorgis A, Lenhard SC, Yerby B, Forsgren MF, Liachenko S, Johansson E, Pilling MA, Peterson RA, Yang X, Williams DP, Ungersma SE, Morgan RE, Brouwer KLR, Jucker BM, Hockings PD. A multi-center preclinical study of gadoxetate DCE-MRI in rats as a biomarker of drug induced inhibition of liver transporter function. *PLoS One* 2018;**13**:e0197213

135. de Greef R, Maloney A, Olsson-Gisleskog P, Schoemaker J, Panagides J. Dopamine D2 occupancy as a biomarker for antipsychotics: quantifying the relationship with efficacy and extrapyramidal symptoms. *Aaps j* 2011;**13**:121-30

136. Nord M, Farde L. Antipsychotic occupancy of dopamine receptors in schizophrenia. *CNS Neurosci Ther* 2011;**17**:97-103

137. Hafezi-Nejad N, Guermazi A, Roemer FW, Eng J, Zikria B, Demehri S. Long term use of analgesics and risk of osteoarthritis progressions and knee replacement: propensity score matched cohort analysis of data from the Osteoarthritis Initiative. *Osteoarthritis Cartilage* 2016;**24**:597-604

138. Reijman M, Bierma-Zeinstra SM, Pols HA, Koes BW, Stricker BH, Hazes JM. Is there an association between the use of different types of nonsteroidal antiinflammatory drugs and radiologic progression of osteoarthritis? The Rotterdam Study. *Arthritis Rheum* 2005;**52**:3137-42

139. Huskisson EC, Berry H, Gishen P, Jubb RW, Whitehead J. Effects of antiinflammatory drugs on the progression of osteoarthritis of the knee. LINK Study Group. Longitudinal Investigation of Nonsteroidal Antiinflammatory Drugs in Knee Osteoarthritis. *J Rheumatol* 1995;**22**:1941-6

140. Tindall EA, Sharp JT, Burr A, Katz TK, Wallemark CB, Verburg K, Lefkowith JB. A 12-month, multicenter, prospective, open-label trial of radiographic analysis of disease progression in osteoarthritis of the knee or hip in patients receiving celecoxib. *Clin Ther* 2002;**24**:2051-63

141. Roemer FW, Hayes CW, Miller CG, Hoover K, Guermazi A. Imaging atlas for eligibility and on-study safety of potential knee adverse events in anti-NGF studies (Part 1). *Osteoarthritis Cartilage* 2015;**23 Suppl 1**:S22-42

142. Roemer FW, Hayes CW, Miller CG, Hoover K, Guermazi A. Imaging atlas for eligibility and on-study safety of potential hip adverse events in anti-NGF studies (Part 2). *Osteoarthritis Cartilage* 2015;**23 Suppl 1**:S43-58

143. Liachenko S, Ramu J, Konak T, Paule MG, Hanig J. Quantitative assessment of MRI T2 response to kainic acid neurotoxicity in rats in vivo. *Toxicol Sci* 2015;**146**:183-91

144. Liachenko S, Ramu J, Paule MG, Hanig J. Comparison of quantitative T(2) and ADC

mapping in the assessment of 3-nitropropionic acid-induced neurotoxicity in rats. *Neurotoxicology* 2018;**65**:52-59

145.   Williams RE, Prior M, Bachelard HS, Waterton JC, Checkley D, Lock EA. MRI studies of the neurotoxic effects of L-2-chloropropionic acid on rat brain. *Magn Reson Imaging* 2001;**19**:133-42

146.   Edmondson DA, Ma RE, Yeh CL, Ward E, Snyder S, Azizi E, Zauber SE, Wells EM, Dydak U. Reversibility of neuroimaging markers influenced by lifetime occupational manganese exposure. *Toxicol Sci* 2019;**172**:181-90

147.   Edmondson DA, Yeh CL, Hélie S, Dydak U. Whole-brain R1 predicts manganese exposure and biological effects in welders. *Arch Toxicol* 2020;**94**:3409-20

148.   Kanda T, Ishii K, Kawaguchi H, Kitajima K, Takenaka D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. *Radiology* 2014;**270**:834-41

149.   Guo BJ, Yang ZL, Zhang LJ. Gadolinium deposition in brain: Current scientific evidence and future perspectives. *Front Mol Neurosci* 2018;**11**:335

150.   Kennan R, Keohane C, Jackson J, Hajdu R, Williams D, Mandala S, Liu H. Non -invasive assay of matrix metalloproteinase induced musculoskeletal syndrome in rats using high resolution magnetic resonance imaging.   International society of magnetic resonance medicine. Seattle, WA2006.

151.   Skeoch S, Weatherley N, Swift AJ, Oldroyd A, Johns C, Hayton C, Giollo A, Wild JM, Waterton JC, Buch M, Linton K, Bruce IN, Leonard C, Bianchi S, Chaudhuri N. Drug-induced interstitial lung disease: A systematic review. *J Clin Med* 2018;**7**

152.   Eaden JA, Skeoch S, Waterton JC, Chaudhuri N, Bianchi SM. How consistently do physicians diagnose and manage drug-induced interstitial lung disease? Two surveys of European ILD specialist physicians. *ERJ Open Res* 2020;**6**

153.   Piga M, Cocco MC, Serra A, Boi F, Loy M, Mariotti S. The usefulness of 99mTc-sestaMIBI thyroid scan in the differential diagnosis and management of amiodarone-induced thyrotoxicosis. *Eur J Endocrinol* 2008;**159**:423-9

154.   Alzahrani AS, Ceresini G, Aldasouqi SA. Role of ultrasonography in the differential diagnosis of thyrotoxicosis: a noninvasive, cost-effective, and widely available but underutilized diagnostic tool. *Endocr Pract* 2012;**18**:567-78

155.   Loy M, Perra E, Melis A, Cianchetti ME, Piga M, Serra A, Pinna G, Mariotti S. Color-flow Doppler sonography in the differential diagnosis and management of amiodarone-induced thyrotoxicosis. *Acta Radiol* 2007;**48**:628-34

156.   Plana JC, Galderisi M, Barac A, Ewer MS, Ky B, Scherrer-Crosbie M, Ganame J, Sebag IA, Agler DA, Badano LP, Banchs J, Cardinale D, Carver J, Cerqueira M, DeCara JM, Edvardsen

T, Flamm SD, Force T, Griffin BP, Jerusalem G, Liu JE, Magalhães A, Marwick T, Sanchez LY, Sicari R, Villarraga HR, Lancellotti P. Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: A report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr* 2014;**27**:911-39

157.   Panday K, Gona A, Humphrey MB. Medication-induced osteoporosis: screening and treatment strategies. *Ther Adv Musculoskelet Dis* 2014;**6**:185-202

158.   Switzer RC, 3rd, Lowry-Franssen C, Benkovic SA. Recommended neuroanatomical sampling practices for comprehensive brain evaluation in nonclinical safety studies. *Toxicol Pathol* 2011;**39**:73-84

159.   Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, Pangalos MN. Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nat Rev Drug Discov* 2014;**13**:419-31

160.   Liachenko S. Translational imaging in toxicology. *Current Opinion in Toxicology* 2020;**23-24**:29-38

161.   Liachenko S, Ramu J. Quantification and reproducibility assessment of the regional brain T(2) relaxation in naïve rats at 7T. *J Magn Reson Imaging* 2017;**45**:700-09

162.   Hanig J, Paule MG, Ramu J, Schmued L, Konak T, Chigurupati S, Slikker W, Jr., Sarkar S, Liachenko S. The use of MRI to assist the section selections for classical pathology assessment of neurotoxicity. *Regul Toxicol Pharmacol* 2014;**70**:641-7

163.   Ramu J, Konak T, Paule MG, Hanig JP, Liachenko S. Longitudinal diffusion tensor imaging of the rat brain after hexachlorophene exposure. *Neurotoxicology* 2016;**56**:225-32

164.   Leptak C, Menetski JP, Wagner JA, Aubrecht J, Brady L, Brumfield M, Chin WW, Hoffmann S, Kelloff G, Lavezzari G, Ranganathan R, Sauer JM, Sistare FD, Zabka T, Wholley D. What evidence do we need for biomarker qualification? *Sci Transl Med* 2017;**9**

165.   Radbruch A, Richter H, Bücker P, Berlandi J, Schänzer A, Deike-Hofmann K, Kleinschnitz C, Schlemmer HP, Forsting M, Paulus W, Martin LF, van Thriel C, Karst U, Jeibmann A. Is small fiber neuropathy induced by gadolinium-based contrast agents? *Invest Radiol* 2020;**55**:473-80

166.   Alkhunizi SM, Fakhoury M, Abou-Kheir W, Lawand N. Gadolinium retention in the central and peripheral nervous system: Implications for pain, cognition, and neurogenesis. *Radiology* 2020;**297**:407-16

167.   McDonald RJ. Assessment of the neurologic effects of intracranial gadolinium deposition using a large population based cohort. Radiological Society of North America 2017.

168.   Forslin Y, Martola J, Bergendal Å, Fredrikson S, Wiberg MK, Granberg T. Gadolinium retention in the brain: An MRI relaxometry study of linear and macrocyclic gadolinium-based contrast agents in multiple sclerosis. *AJNR Am J Neuroradiol* 2019;**40**:1265-73

169. Parillo M, Sapienza M, Arpaia F, Magnani F, Mallio CA, D'Alessio P, Quattrocchi CC. A structured survey on adverse events occurring within 24 hours after intravenous exposure to gadodiamide or gadoterate meglumine: A controlled prospective comparison study. *Invest Radiol* 2019;**54**:191-97

170. Arnold ME, Booth B, King L, Ray C. Workshop Report: Crystal City VI-Bioanalytical method validation for biomarkers. *Aaps j* 2016;**18**:1366-72

171. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, Boellaard R, Bohndiek SE, Brady M, Brown G, Buckley DL, Chenevert TL, Clarke LP, Collette S, Cook GJ, deSouza NM, Dickson JC, Dive C, Evelhoch JL, Faivre-Finn C, Gallagher FA, Gilbert FJ, Gillies RJ, Goh V, Griffiths JR, Groves AM, Halligan S, Harris AL, Hawkes DJ, Hoekstra OS, Huang EP, Hutton BF, Jackson EF, Jayson GC, Jones A, Koh DM, Lacombe D, Lambin P, Lassau N, Leach MO, Lee TY, Leen EL, Lewis JS, Liu Y, Lythgoe MF, Manoharan P, Maxwell RJ, Miles KA, Morgan B, Morris S, Ng T, Padhani AR, Parker GJ, Partridge M, Pathak AP, Peet AC, Punwani S, Reynolds AR, Robinson SP, Shankar LK, Sharma RA, Soloviev D, Stroobants S, Sullivan DC, Taylor SA, Tofts PS, Tozer GM, van Herk M, Walker-Samuel S, Wason J, Williams KJ, Workman P, Yankeelov TE, Brindle KM, McShane LM, Jackson A, Waterton JC. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 2017;**14**:169-86

172. Willmann JK, van Bruggen N, Dinkelborg LM, Gambhir SS. Molecular imaging in drug development. *Nat Rev Drug Discov* 2008;**7**:591-607

173. Morgan P, Van Der Graaf PH, Arrowsmith J, Feltner DE, Drummond KS, Wegner CD, Street SD. Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. *Drug Discov Today* 2012;**17**:419-24

174. Nickolls SA, Gurrell R, van Amerongen G, Kammonen J, Cao L, Brown AR, Stead C, Mead A, Watson C, Hsu C, Owen RM, Pike A, Fish RL, Chen L, Qiu R, Morris ED, Feng G, Whitlock M, Gorman D, van Gerven J, Reynolds DS, Dua P, Butt RP. Pharmacology in translation: the preclinical and early clinical profile of the novel α2/3 functionally selective GABA(A) receptor positive allosteric modulator PF-06372865. *Br J Pharmacol* 2018;**175**:708-25

175. Awadalla M, Hassan MZO, Alvi RM, Neilan TG. Advanced imaging modalities to detect cardiotoxicity. *Curr Probl Cancer* 2018;**42**:386-96

176. Michel L, Rassaf T, Totzeck M. Biomarkers for the detection of apparent and subclinical cancer therapy-related cardiotoxicity. *J Thorac Dis* 2018;**10**:S4282-s95

177. Cardinale D, Colombo A, Torrisi R, Sandri MT, Civelli M, Salvatici M, Lamantia G, Colombo N, Cortinovis S, Dessanai MA, Nolè F, Veglia F, Cipolla CM. Trastuzumab-induced cardiotoxicity: clinical and prognostic implications of troponin I evaluation. *J Clin Oncol* 2010;**28**:3910-6

178.   Saba L, Biswas M, Kuppili V, Cuadrado Godia E, Suri HS, Edla DR, Omerzu T, Laird JR, Khanna NN, Mavrogeni S, Protogerou A, Sfikakis PP, Viswanathan V, Kitas GD, Nicolaides A, Gupta A, Suri JS. The present and future of deep learning in radiology. *Eur J Radiol* 2019;**114**:14-24

179.   Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;**5**:4006

180.   Rohé M, Oubel E, Grisoni M, Djedjos S, Myers R, Loomba H, Middleto M. Feasibility of using deep learning techniques to assess hepatic fibrosis directly from magnetic resonance elastography source images. The Liver Meeting® 2018, AASLD. San Francisco, CA, USA2018.

181.   Lee JI, Lee HW, Kim SU, Ahn SH, Lee KS. Follow-up Liver stiffness measurements after liver resection influence oncologic outcomes of hepatitis-B-associated hepatocellular carcinoma with liver cirrhosis. *Cancers (Basel)* 2019;**11**

182.   Rexhepaj E, Ramos C, Liu Y, Huet B, Lucidarme O, Boujemaa N. High risk fibrosis score prediction using computed tomography imaging. The digital international liver congress, EASLD. virtual2020.

183.   Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, Bosch TC, Cryan JF, Gilbert SF, Goodnight CJ, Lloyd EA, Sapp J, Vandenkoornhuyse P, Zilber-Rosenberg I, Rosenberg E, Bordenstein SR. Getting the hologenome concept right: An eco-evolutionary framework for hosts and their microbiomes. *mSystems* 2016;**1**

184.   Rosenberg E, Zilber-Rosenberg I. Microbes drive evolution of animals and plants: The hologenome concept. *mBio* 2016;**7**:e01395

185.   Gosset G. Production of aromatic compounds in bacteria. *Curr Opin Biotechnol* 2009;**20**:651-8

186.   Ackermann W, Coenen M, Schrödl W, Shehata AA, Krüger M. The influence of glyphosate on the microbiota and production of botulinum neurotoxin during ruminal fermentation. *Curr Microbiol* 2015;**70**:374-82

187.   Shehata AA, Schrödl W, Aldin AA, Hafez HM, Krüger M. The effect of glyphosate on potential pathogens and beneficial members of poultry microbiota in vitro. *Curr Microbiol* 2013;**66**:350-8

188.   Nielsen LN, Roager HM, Casas ME, Frandsen HL, Gosewinkel U, Bester K, Licht TR, Hendriksen NB, Bahl MI. Glyphosate has limited short-term effects on commensal bacterial community composition in the gut environment due to sufficient aromatic amino acid levels. *Environ Pollut* 2018;**233**:364-76

189.   Zucko J, Dunlap WC, Shick JM, Cullum J, Cercelet F, Amin B, Hammen L, Lau T,

Williams J, Hranueli D, Long PF. Global genome analysis of the shikimic acid pathway reveals greater gene loss in host-associated than in free-living bacteria. *BMC Genomics* 2010;**11**:628

190. Tomova A, Bukovsky I, Rembert E, Yonas W, Alwarith J, Barnard ND, Kahleova H. The effects of vegetarian and vegan diets on gut microbiota. *Front Nutr* 2019;**6**:47

191. National Academies of Sciences E, Medicine, Division on E, Life S, Board on Life S, Board on Environmental S, Toxicology, Committee on advancing understanding of the implications of environmental-chemical interactions with the human M. The national academies collection: Reports funded by National Institutes of Health. Environmental chemicals, the human microbiome, and health risk: A research strategy. Washington (DC): National Academies Press (US)

Copyright 2018 by the National Academy of Sciences. All rights reserved., 2017

192. Chassaing B, Koren O, Goodrich JK, Poole AC, Srinivasan S, Ley RE, Gewirtz AT. Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature* 2015;**519**:92-6

193. Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* 2008;**3**:213-23

194. Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, Israeli D, Zmora N, Gilad S, Weinberger A, Kuperman Y, Harmelin A, Kolodkin-Gal I, Shapiro H, Halpern Z, Segal E, Elinav E. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* 2014;**514**:181-6

195. Requena T, Martínez-Cuesta MC, Peláez C. Diet and microbiota linked in health and disease. *Food Funct* 2018;**9**:688-704

196. Deehan EC, Duar RM, Armet AM, Perez-Muñoz ME, Jin M, Walter J. Modulation of the gastrointestinal microbiome with nondigestible fermentable carbohydrates to improve human health. *Microbiol Spectr* 2017;**5**

197. Csáki FK, Sebestyén É. Who will carry out the tests that would be necessary for proper safety evaluation of food emulsifiers? *Food Science and Human Wellness* 2019;**8**:126-35

198. Ruiz-Ojeda FJ, Plaza-Díaz J, Sáez-Lara MJ, Gil A. Effects of sweeteners on the gut microbiota: A review of experimental studies and clinical trials. *Adv Nutr* 2019;**10**:S31-s48

199. Rojo D, Méndez-García C, Raczkowska BA, Bargiela R, Moya A, Ferrer M, Barbas C. Exploring the human microbiome from multiple perspectives: factors altering its composition and function. *FEMS Microbiol Rev* 2017;**41**:453-78

200. Aura A-M, Maukonen J. One Compartment Fermentation Model. In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on health: In vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 281-93

201. Barroso E, Cueva C, Peláez C, Martínez- Cuesta M, Requena T. The computer-controlled multicompartmental dynamic model of the gastrointestinal system SIMGI In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on health: in vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 319-29

202. Minekus M. The TNO Gastro-Intestinal Model (TIM) In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on health: in vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 37-47

203. Thuenemann E. Dynamic Digestion Models: General Introduction In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on health: in vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 33-37

204. Thuenemann E, Mandalari G, Gillian T. Rich G, Faulks R. Dynamic Gastric Model (DGM). In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on health: in vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 47-61

205. Van de Wiele T, Van den Abbeele P, Ossieur W, Possemiers S, Marzorati M. The Simulator of the Human Intestinal Microbial Ecosystem (SHIME®) In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on health: in vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 305-19

206. Venema K. The TNO In Vitro Model of the Colon (TIM-2) In: Verhoeckx K, Cotter P, López-Expósito I, Kleiveland C, Lea T, Mackie A, Requena T, Swiatecka D, Wichers H, editors. The impact of food bioactives on Health: in vitro and ex vivo models. Cham (CH): Springer Copyright 2015, The Editor(s) (if applicable) and the Author(s). 2015. pp. 293-305

207. Barroso E, Cueva C, Peláez C, Martínez-Cuesta MC, Requena T. Development of human colonic microbiota in the computer-controlled dynamic SIMulator of the gastroIntestinal tract SIMGI. *LWT - Food Science and Technology* 2015;**61**:283-89

208. Bein A, Shin W, Jalili-Firoozinezhad S, Park MH, Sontheimer-Phelps A, Tovaglieri A, Chalkiadaki A, Kim HJ, Ingber DE. Microfluidic organ-on-a-chip models of human intestine. *Cell Mol Gastroenterol Hepatol* 2018;**5**:659-68

209. Clavel T, Lagkouvardos I, Blaut M, Stecher B. The mouse gut microbiome revisited: From complex diversity to model ecosystems. *Int J Med Microbiol* 2016;**306**:316-27