Workshop Report

Trust Your Gut: Establishing Confidence in Gastrointestinal Models – An Overview of the State of the Science and Contexts of Use

Susan Debad[1], David Allen[2], Omari Bandele[3], Colin Bishop[4], Michaela Blaylock[2], Paul Brown[5], Maureen K. Bunger[6], Julia Y. Co[7], Lynn Crosby[3], Amber B. Daniel[2], Steve S. Ferguson[8], Kevin Ford[9], Gonçalo Gamboa da Costa[10], Kristin H. Gilchrist[11], Matthew W. Grogg[12], Maureen Gwinn[13], Thomas Hartung[14], Simon P. Hogan[15], Ye Eun Jeong[16], George E. N. Kass[17], Elaina Kenyon[13], Nicole C. Kleinstreuer[18], Ville Kujala[19], Patrik Lundquist[20], Joanna Matheson[21], Shaun D. McCullough[22], Angela Melton-Celsa[23], Steven Musser[3], Ilung Oh[24], Oluwakemi B. Oyetade[2], Sarita U. Patil[25], Elijah J. Petersen[26], Nakissa Sadrieh[3], Christie M. Sayes[27], Benjamin S. Scruggs[6], Yu-Mei Tan[28], Bill Thelin[6], M. Tyler Nelson[12], José V. Tarazona[29], John F. Wambaugh[13], Jun-Young Yang[24], Changwoo Yu[24] and Suzanne Fitzpatrick[3]

[1]SJD Consulting LLC, Ijamsville, MD, USA; [2]Inotiv, Inc., RTP, NC, USA; [3]Center for Food Safety and Applied Nutrition, College Park, MD, USA; [4]Wake Forest Institute for Regenerative Medicine, Winston Salem, NC, USA; [5]Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA; [6]Altis Biosystems, Inc., Research Triangle Park, NC, USA; [7]Complex in vitro Systems, Safety Assessment, Genentech, South San Francisco, CA, USA; [8]Mechanistic Toxicology Branch, Division of Translational Toxicology, NIEHS, National Institutes of Health, Bethesda, MD, USA; [9]Office of Clinical Pharmacology, US Food and Drug Administration, Silver Spring, MD, USA; [10]FDA National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, USA; [11]4D Bio^3 Center for Biotechnology, Uniformed Services University, Bethesda, MD, USA; [12]711th Human Performance Wing, Air Force Research Laboratory, Wright-Patterson AFB, OH, USA; [13]U.S. Environmental Protection Agency, Office of Research and Development, RTP, NC, USA; [14]Johns Hopkins University, Bloomberg School of Public Health and Whiting School of Engineering, Doerenkamp-Zbinden-Chair for Evidence-based Toxicology, Center for Alternatives to Animal Testing (CAAT), Baltimore, MD, USA and University of Konstanz, CAAT-Europe, Konstanz, Germany; [15]Mary H. Weiser Food Allergy Center, Department of Pathology, Michigan Medicine, University of Michigan, Ann Arbor, MI, USA; [16]Division of Applied Regulatory Science, Office of Clinical Pharmacology, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA; [17]European Food Safety Authority, Parma, Italy; [18]NIEHS/NICEATM, RTP, NC, USA; [19]Emulate Inc., Boston, MA, USA; [20]Department of Pharmacy, Uppsala University, Uppsala, Sweden; [21]U.S. Consumer Product Safety Commission, Rockville, MD, USA; [22]Exposure and Protection, RTI International, Durham, NC, USA; [23]Department of Microbiology and Immunology, Uniformed Services University, Bethesda, MD, USA; [24]Toxicology Research Division, Ministry of Food and Drug Safety, National Institute of Food and Drug Safety Evaluation, Cheongju, Republic of Korea; [25]Divisions of Allergy and Clinical Immunology, Departments of Medicine and Pediatrics, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [26]Biosystems and Biomaterials Division, National Institute of Standards and Technology, Gaithersburg, MD, USA; [27]Department of Environmental Science, Baylor University, Waco, TX, USA; [28]U.S. Environmental Protection Agency, Office of Pesticide Programs, Research Triangle Park, NC, USA; [29]Spanish National Environmental Health Centre, Instituto de Salud Carlos III, Madrid, Spain

Received March 26, 2024;
Epub May 15, 2024;

Correspondence:
Thomas Hartung, MD, PhD,
Center for Alternatives to Animal Testing (CAAT),
Johns Hopkins University,
615 N Wolfe St., Baltimore, MD, 21205, USA
(THartun1@jhu.edu)

ALTEX 41(3), 402-424.
doi:10.14573/altex.2403261

Abstract

The webinar series and workshop titled "Trust Your Gut: Establishing Confidence in Gastrointestinal Models – An Overview of the State of the Science and Contexts of Use" was co-organized by NICEATM, NIEHS, FDA, EPA, CPSC, DoD, and the Johns Hopkins Center for Alternatives to Animal Testing (CAAT) and hosted at the National Institutes of Health in Bethesda, MD, USA on October 11-12, 2023. New approach methods (NAMs) for assessing issues of gastrointestinal tract (GIT)-related toxicity offer promise in addressing some of the limitations associated with animal-based assessments. GIT NAMs vary in complexity, from two-dimensional monolayer cell line-based systems to sophisticated 3-dimensional organoid systems derived from human primary cells. Despite advances in GIT NAMs, challenges remain in fully replicating the complex interactions and processes occurring within the human GIT. Presentations and discussions addressed regulatory needs, challenges, and innovations in incorporating NAMs into risk assessment frameworks; explored the state of the science in using NAMs for evaluating systemic toxicity, understanding absorption and pharmacokinetics, evaluating GIT toxicity, and assessing potential allergenicity; and discussed strengths, limitations, and data gaps of GIT NAMs as well as steps needed to establish confidence in these models for use in the regulatory setting.

Plain language summary

Non-animal methods to assess whether chemicals may be toxic to the human digestive tract promise to complement or improve on animal-based methods. These approaches, which are based on human or animal cells and/or computer models, are faced with their own technical challenges and need to be shown to predict adverse effects in humans. Regulators are tasked with evaluating submitted data to best protect human health and the environment. A webinar series and workshop brought together scientists from academia, industry, military, and regulatory authorities from different countries to discuss how non-animal methods can be integrated into the risk assessment of drugs, food additives, dietary supplements, pesticides, and industrial chemicals for gastrointestinal toxicity.

1 Introduction

1.1 Objectives of meeting/background

One of the most common chemical exposure routes is via the gastrointestinal tract (GIT). Most orally administered substances, including drugs, food additives, and dietary supplements, as well as pesticides and various industrial chemicals, may be initially metabolized within the GIT; and both the parent substance and any formed metabolites may have direct effects on and/or be absorbed through the GIT. During these processes, either parent substances, metabolites, or both may damage the GIT, leading to local or systemic toxicity or allergenicity. Thus, it is crucial to conduct screening for potential toxicants targeting the GIT. Even in the absence of GIT toxicity, understanding the extent of absorption is key to gaining insights into systemic exposure levels, offering valuable information about potential toxicity in other target organs.

While *in vivo* animal studies continue to be used to understand GIT absorption and toxicity because they recapitulate many human features in an integrated way, animal models have anatomical, physiological, and biochemical differences from humans. Unsurprisingly, this highlights the concept that human cells likely represent the best model for human toxicology. Mammalian cell culture methods of varying complexity, up to and including 3-dimensional models, may be able to address some of the limitations associated with existing safety assessment approaches. These methods have a variety of applications that can be broadly useful for researchers, pharmaceutical or chemical manufacturers, and regulators.

Accordingly, NICEATM, NIEHS, FDA, EPA, CPSC, DoD, and CAAT organized a workshop on the state of the science for using new approach methods (NAMs) to predict GIT absorption and toxicity. The workshop centered around three main potential uses/functions of NAMs:

– Evaluating systemic toxicity and understanding absorption/pharmacokinetics;
– Predicting GIT toxicity; and
– Evaluating potential allergenicity.

1.1.1 Evaluating systemic toxicity, understanding absorption/pharmacokinetics

In the context of pharmaceutical development, chemical production, or regulatory assessments, de-risking is a broad term that refers to the strategies and measures implemented to proactively identify, manage, and reduce the risk associated with a chemical, substance, or drug. These risks are understood as a function of both the hazard posed by the substance and the potential exposure to that substance. Thus, de-risking not only implies lowering (but not eliminating) risk through the identification of hazards that can be assessed with existing methods, but also through evaluating these hazards in conjunction with exposure data to inform a comprehensive dose-response assessment.

De-risking systemic toxicity through hazard evaluation, a standard, integral component of the overall de-risking process for substances that could potentially enter systemic circulation – including virtually all orally administered drugs, food additives, dietary supplements, pesticides, food contaminants, and various industrial chemicals that may contact the GIT – is typically accomplished through a combination of *in vitro* tests, animal models, and computational predictions. These tools are used to assess a substance's toxicological profile[1], integrating both hazard identification and exposure considerations to provide a holistic understanding of potential exposure risks. Traditional *in vivo* methods involve dose-response studies and repeated exposure tests to determine safety margins and identify potential systemic effects. Hazard evaluation is intricately linked with the pharmacokinetic profile of a substance, which includes four key processes: absorption, distribution, metabolism, and excretion (ADME). The ADME profile determines the actual exposure of internal organ systems to the substance, which in turn influences its potential to cause systemic toxicity. For pharmaceuticals, ADME profiling has been critical for reducing drug failures and demonstrates how critical ADME can be for initial assessment (Tsaioun et al., 2016). If a substance is not absorbed or is absorbed at low levels, it is less likely to reach meaningful systemic circulation and therefore less likely to cause adverse effects in organs or tissues distant from the site of exposure. Consequently,

[1] See, for example, guidance documents from U.S. FDA, https://www.fda.gov/media/152544/download and U.S. EPA, https://www.regulations.gov/document/EPA-HQ-OPPT-2016-0654-0108

a comprehensive assessment of a substance's absorption through the GIT can be indispensable for predicting its safety profile and could preclude the need for systemic studies.

1.1.2 Predicting gastrointestinal toxicity

Gastrointestinal toxicity encompasses the adverse effects that pharmaceuticals, food additives, dietary supplements, pesticides, and various other chemicals can have directly on the digestive system. These effects may range from mild discomfort, such as nausea and diarrhea, to severe complications such as ulcers, bleeding, or even perforation of the GIT (Gelberg, 2018). In both pharmaceutical development and the regulatory testing of consumer products, assessing GIT toxicity is of paramount importance because the digestive system is the first point of contact for ingested substances, making it especially susceptible to their effects. Additionally, the complex ecosystem of the GIT, with its diverse microbiota, is highly sensitive to disturbances from exogenous compounds, which can lead to dysbiosis and a spectrum of long-term health issues (Chiu et al., 2020; Calero-Medina et al., 2023). Consequently, evaluating GIT toxicity is a stringent requirement of regulatory bodies worldwide, to ensure comprehensive consumer protection and maintain public health standards. *In vitro* assessment of GIT toxicity has utilized transformed cell lines *in vitro* (e.g., Caco-2 cells (Artursson et al., 2001)), as well as biochemical assays. *In vitro* techniques, along with animal studies, are designed to uncover a range of GIT reactions, from changes in the mucosal lining to alterations in motility and absorption.

1.1.3 Evaluating potential allergenicity

Evaluation of potential allergenicity is a fundamental concern across various industries, as the introduction of new chemicals into the market – whether drugs, cosmetics, food additives, or industrial agents – poses the risk of triggering allergic reactions in sensitized end-users. The consequences of such reactions can be significant, encompassing everything from mild discomfort to life-threatening anaphylaxis, making allergenicity assessment a critical step in the regulatory approval process. Traditional risk assessment for allergenicity, particularly in the context of food allergens and new pharmaceuticals, has relied on a combination of *in vitro* assays (e.g., IgE binding tests, ELISAs, *in vitro* skin sensitization tests), animal models (e.g., *in vivo* skin sensitization tests), and clinical data (e.g., skin-prick tests) (Crevel et al., 2014; Pali-Schöll et al., 2019).

1.2 Limitations of traditional approaches necessitate new approach methods

While human clinical trials are used to assess safety and efficacy of therapeutic chemicals, animal studies have long been considered the "gold standard" for comprehensive toxicological assessment of consumer products, that is, identification of chemical hazard and dose-response assessment. While basic *in vitro* techniques for assessing toxicity and allergenicity can be valuable, particularly for initial toxicity screening, these methods often lack the complexity to fully mimic the intricate biological interactions occurring within an organism. Nevertheless, many *in vitro* assays are currently accepted by regulatory bodies to support chemical safety assessment for a number of endpoints, such as ocular toxicity, skin sensitization, and phototoxicity[2]. However, animal models also have notable limitations. Rats, for instance, have no gallbladders and, unlike humans, are obligate nose-breathers; furthermore, they have no emetic response. Thus, animal responses may not accurately mirror human physiology. For example, while rodent studies have historically been used to understand GIT absorption and toxicity, this model has well-documented anatomical, physiological, and biochemical differences from humans, including the presence of a forestomach in rats and mice, differences in ADME processes, diet, metabolism, microbial flora, and immune responses (including adaptive immune receptor repertoires) (Chi et al., 1982; Mestas and Hughes, 2004; Van Norman, 2019). Furthermore, in many studies, dosage and mode of administration (e.g., suspension of test article in corn oil with administration by gavage) differ from human exposures. Species-specific differences can result in toxicity and efficacy outcomes that may not accurately reflect the human situation (Atkins et al., 2020; Marshall et al., 2023). Additionally, although *in vivo* toxicity studies primarily use outbred animals, which provide some genetic variability, they fail to fully recapitulate the genetic variability within human populations. Genetic variability can influence both toxicity and efficacy of pharmaceuticals (Ventola, 2013; Pirmohamed, 2023). Further, most studies use healthy animals, which fail to account for the comorbidities often present in the human population.

Higher-order animals like non-human primates are reasonably good at predicting human GIT toxicity of drugs (Monticello et al., 2017). However, there are challenges with using such models, including cost and availability of animals. Consequently, drugs with GIT-adverse effects may progress through early stages of drug development into clinical trials, in which GIT disorders are the most common category of adverse events (Federer et al., 2016). Often, pharmaceuticals exhibiting GIT toxicity in clinical trials may have shown GIT toxicity in preclinical animal studies, but progression into trials occurred because the toxicity was considered non-serious, monitorable, and acceptable given the potential benefit. In some instances, GIT toxicity in clinical trials or during marketing is not adequately balanced by the benefit, and use of the drug may be suspended (Qureshi et al., 2011; Cook et al., 2014). GIT side effects of approved drugs can be dose limiting, especially for chemotherapeutics, and can reduce patient compliance (O'Reilly et al., 2020). *In vitro* assays capable of screening for drug-induced human GIT toxicity are likely to add value to the overall drug discovery and development process (Peters et al., 2020).

Despite advances in predictive assays for allergenicity, the intricate nature of human immune responses presents a substantial re-

[2] See, for example, https://www.epa.gov/system/files/documents/2024-01/oppt-ncd-eye-irritation-framework-frn-final-12-13-2023.pdf; https://www.oecd.org/env/test-no-442e-in-vitro-skin-sensitisation-9789264264359-en.htm; https://www.oecd-ilibrary.org/environment/test-no-498-in-vitro-phototoxicity-reconstructed-human-epidermis-phototoxicity-test-method_7b2f9ea0-en

search gap. *In vivo* allergenicity testing faces significant limitations due to ethical considerations and the constraints of animal welfare regulations. Species-specific differences in immune responses can lead to unreliable extrapolation of results from animals to humans, and sometimes even between humans, highlighting the importance of uncertainty considerations. Although animal models have provided insight into skin sensitization and the local lymph node assay was shown to be predictive for known respiratory sensitizers (Dearman et al., 2013), the predictive power of animal assays often falls short for food allergenicity (Bøgh et al., 2016), for which the biological mechanisms are complex and incompletely understood. Notably, validated and predicted models are absent (Kazemi et al., 2023) and existing models are profoundly limited, with many known human allergens failing to provoke a comparable response in animal tests, potentially resulting in underestimation of risks or failure to identify novel allergenic compounds.

Both regulatory agencies and industry stakeholders are invested in improving the methodologies for assessing GIT toxicity and allergenicity through the use of novel assays that may better recapitulate human reactions and thus more accurately predict adverse effects without using animals. Such approaches are collectively called new approach methods (NAMs) (Kavlock et al., 2018; Kleinstreuer et al., 2018). While use of the term can vary, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) defines NAMs as "*...any technology, methodology, approach, or combination thereof that can be used to provide information on chemical hazard and risk assessment and supports replacement, reduction, or refinement of animal use (3Rs)*"[3]. As detailed in the remainder of this report, NAMs provide unique opportunities to assess absorption, GIT toxicity, adverse effects on the microbiome, or allergenicity resulting from orally administered substances without animal testing. NAMs will allow incorporation of human genetic diversity in ADME and toxicity assessments by sourcing cells from multiple individuals. In addition, it may be increasingly possible to use *in vitro* models of disease or incorporate other factors into NAMs to address the impact of human health status or other environmental factors on the response to drugs or other chemicals.

Moreover, NAMs have the potential to reduce – and in certain circumstances potentially replace – the need for existing animal testing approaches, thus aligning more closely with the principles of the 3Rs (replacement, reduction, and refinement) that support the humane use of animals in scientific research (Fenwick et al., 2009). Any progress in NAMs, therefore, marks a significant stride in both ethical research practices and in the pursuit of more accurate, expedited human toxicological assessments. While regulatory agencies and industry stakeholders worldwide are increasingly focused on refining NAMs and integrating them into risk assessment frameworks, the promise of NAMs is contingent on their validation and regulatory acceptance, and many hurdles remain. The next section will discuss the characteristics required to establish confidence in NAMs and the associated challenges of incorporating these methods into the regulatory landscape. Progress of agencies to date as well as future plans for continued NAM integration will be described, as will strategic considerations to assist the transition, including communication, education, and collaboration.

2 Integrating NAMs into the regulatory landscape for enhanced safety assessments: Needs, challenges, and innovations

From the perspectives of their various agencies, four webinar speakers set the stage for the in-person workshop by describing efforts and challenges to integrating NAMs into a regulatory framework. The speakers were Maureen R. Gwinn, Principal Deputy Assistant Administrator for Research and Development at the U.S. EPA's Office of Research and Development; Steven Musser, Deputy Center Director for Scientific Operations at the U.S. Food and Drug Administration (FDA); George Kass, Lead Expert in Toxicology at the European Food Safety Authority (EFSA); and José V. Tarazona, Research Professor and Head Risk Assessment Unit, Spanish National Environmental Health Centre, Instituto de Salud Carlos III.

2.1 Current landscape: Regulatory agency needs and challenges

The regulatory agencies that presented at this workshop oversee a broad range of products to ensure public safety and environmental protection. Products with GIT relevance include pharmaceuticals, food additives, pesticides, industrial chemicals, and cosmetics (when accidentally ingested). While each agency has its unique focus and jurisdiction, traditional toxicity testing has a set of standard, globally recognized methodologies that are often shared among agencies. As regulatory agencies worldwide attempt to improve the human relevance of traditional methods by integrating NAMs into their risk assessment frameworks, a common set of needs and challenges is emerging – primarily around determining which NAMs might be applicable to regulatory decision making.

2.1.1 Scientific confidence

As regulatory agencies are anchored to the principle of safeguarding public and environmental health, any methodologies employed – traditional or novel – must undergo rigorous evaluation to ensure scientific confidence. This stringent scrutiny ensures that NAMs can reliably replicate and, ideally, improve upon the outcomes derived from traditional tests, and offer a sound scientific basis for regulatory decision making. Webinar speakers noted four main components central to building the scientific confidence needed to successfully incorporate NAMs into regulatory frameworks: technical validation, standardization, transparency, and biological relevance.

Technical validation, standardization, and transparency

The Organisation for Economic Co-operation and Development (OECD) defines test method validation as "*a process based on sci-*

[3] https://ntp.niehs.nih.gov/whatwestudy/niceatm/resources-for-test-method-developers/submissions

entifically sound principles by which the reliability and relevance of a particular test, approach, method, or process are established for a specific purpose."[4] Validation involves rigorous assessment of a method's accuracy, reliability, reproducibility, and relevance across laboratories and under various conditions, and can involve inter-laboratory studies, known as round-robin testing. ICCVAM states further that the processes used for validation "*should allow for efficient and timely development of NAMs that are reliable, fit-for-purpose, and provide information relevant to the species of interest*."[5]

Typically, new assays are validated by comparing their results with outcomes obtained from established and accepted testing methods. Assessing NAMs in this way can demonstrate their ability to reliably replicate known effects or accurately predict outcomes that would be observed in a more traditional assessment framework. Thus, validation is particularly challenging for data-poor chemicals and for NAMs measuring endpoints not covered by typical animal study protocols, for which insufficient or no relevant scientific data are available. In the absence of a reliable benchmark, determining the relevance of NAM results is difficult. However, "*In some instances, the NAM may provide biologically relevant information, mechanistic insights, or sufficiently sensitive endpoints that are adequate for the regulatory decision-making process, and a comparison to data from traditional animal test methods may not be necessary*."[5]

Standardization, another important characteristic of any mature assay, involves the development of detailed protocols, reference compounds, and quality control procedures that support an assay being consistently performed within and across laboratories. For some assays, acceptance criteria and performance standards can be used by laboratories to confirm that they can perform a method sufficiently well. When possible, only readily available reagents with consistent characteristics among batches and suppliers should be used in standardized methods. Standardization underlies confidence in data and allows data to be compared across studies and regulatory submissions.

Finally, transparency underpins stakeholder confidence in NAMs. Clear and careful documentation is a key component of transparency – not just documentation of protocols and procedures, but of the scientific rationale, the uncertainties, and the limitations intrinsic to these novel methods. Per ICCVAM, "*Transparency facilitates trust in the use of NAMs and thereby hastens the pace of an agency's regulatory decision-making process and potential regulatory acceptance or qualification. A NAM's relevance to the species, COU [context of use], and technical characterization should be transparently communicated to peer reviewers, the scientific community, and to the public*."[5] Understanding the limitations of each NAM is particularly crucial as it helps delineate the specific contexts in which these methods can be considered dependable. Transparency necessitates the open sharing of methods, data, and results to allow for thorough peer review and scrutiny by the broader scientific community. Peer review also acts as a cornerstone of validation, contributing to the collective confidence in innovative approaches.

Human biological relevance

While technically an aspect of validation, human biological relevance is a particularly important aspect of NAMs that are to be used for risk assessment measures. According to ICCVAM, "*The relevance of a NAM describes the relationship between the test and the effect in the target species and whether the test method is meaningful and useful for a defined purpose, with the limitations identified. Adequate demonstration of the relevance of a NAM is an important contributor to confidence in a NAM*."[5] Specifically, biological response is the bridge that connects technical measurements with meaningful implications, ensuring that assays are relevant to human health. To be useful in the regulatory setting, NAMs must not only correlate with known outcomes from traditional tests or *in vivo* human data but also ideally predict toxicological effects and other relevant endpoints. An important consideration is whether *in vitro* test doses align with realistic human exposure levels.

The demand for human biological relevance calls for an in-depth assessment of a NAM's biological pertinence. For instance, models aiming to represent human GIT physiology should recapitulate both the structure and the function of the human GIT as accurately as possible. Depending on COU, this may involve moving beyond a singular focus on intestinal epithelia to include aspects such as mucus production, Peyer's patches, dynamic flow, peristaltic movement, and the human microbiota, for example – features that may be essential to truly understand a substance's metabolism, absorption, toxicity, and subsequent impacts on human health (Blutt et al., 2017).

The challenge of recapitulating human variability and complexity is non-trivial (Jamei et al., 2009). *In vivo* models often fail to capture the full spectrum of human responses to substances and, ideally, NAMs would exhibit improved predictive outcomes across various demographics, incorporating differences in genetic makeup, age, and health status. A consideration of health status is paramount, since traditional models generally assume a baseline of perfect health and are far from representative of ethnic variability, chronic illnesses, or other conditions affecting a large percentage of humans. Assessing the impact of substances on particularly vulnerable groups, such as infants, young children, or the elderly, is a critical regulatory consideration.

Both traditional animal models and NAMs have inherent limitations, and thus some form of extrapolation is necessary with either approach. Animal models require interspecies extrapolation, while *in vitro* models need *in-vitro*-to-*in-vivo* extrapolation (IVIVE) (Rotroff et al., 2010; Coecke et al., 2013; Bell et al., 2018). In essence, even sophisticated NAMs are simplified representations of the complex human organism, and they may not be able to replicate intricate interactions between body systems or the long-term development of adverse outcomes. Thus, human relevance may

[4] https://one.oecd.org/document/ENV/JM/MONO(2005)14/en/pdf

[5] https://ntp.niehs.nih.gov/sites/default/files/2024-03/VWG_Report_27Feb2024_FD_508.pdf

ultimately also involve modeling complex pathways between interacting systems, such as the GIT, brain, lungs, and cardiovascular system, using physiologically based pharmacokinetic (PBPK) or other physiological models.

2.1.2 Efficiency and flexibility

In the risk assessment setting, assay efficiency and flexibility are paramount. While invaluable for historic insights into toxicological effects, animal studies are often time-consuming and resource-intensive and may not provide human-relevant information. To be of use to regulators, NAMs would ideally exhibit improved efficiency for assessing absorption and bioavailability, for example. More efficient methods could expedite the entire risk assessment process, possibly decreasing reliance on extensive *in vivo* studies and eliminating unnecessary repetitive testing often associated with traditional approaches – ultimately saving time and resources without compromising scientific integrity or public safety.

The portfolios of most regulatory agencies are vast, diverse, and rapidly changing. The U.S. FDA is a case in point, with a current regulatory portfolio including substances used in novel food sources (e.g., lab-grown meat, insects, or fungi), novel food components (e.g., pea or potato protein), novel food technologies (e.g., GE and GMO crops, nanotechnologies used in foods or packaging), and dietary supplements (e.g., live microbials). Flexibility describes the ability of the regulatory process to cope with this diversity, potentially necessitating assays (or a range of assays) applicable to multiple decision contexts, including evolving threats and challenges (e.g., the effects of chemicals on pollinators or the safety of nanomaterials).

Flexibility could also extend to the capacity to assess complex chemical mixtures often encountered through foods, environmental exposures, and pharmaceuticals. Chemicals can interact synergistically or antagonistically[6] (Martin, 2023) or may not interact at all. Exposures can also vary considerably as chemicals and products break down once in the environment. Further, assessing risk from multiple chemicals that share the same mode of action or cause the same type of effect is a significant challenge, particularly when chemical exposures may come from various sources and/or regulatory jurisdictions. Analyzing complex mixtures requires advanced methods that can accurately identify and quantify each constituent, which can be technically challenging and resource-intensive. The relatively low throughput of many current models makes them ill-equipped to handle the large number of possible chemical exposure combinations. As a result, NAMs may provide a future opportunity for regulatory agencies to advance their evaluation of multiple chemical exposures by accounting for such complexities, ensuring that the efficiency of the assessment process does not come at the expense of thoroughness and reliability.

2.2 Innovations and current progress on integration of NAMs into the regulatory space

As regulatory agencies worldwide continue to evolve in response to scientific advances and societal expectations, the integration of NAMs into the regulatory landscape marks a significant paradigm shift (Collins et al., 2008; Thomas et al., 2013; Cote et al., 2016; Kavlock et al., 2018; Krewski et al., 2020; Pallocca et al., 2022; Chang et al., 2022; Petersen, 2022; Turner et al., 2023)[7]. This section explores the progress made by the four regulatory agencies represented in the pre-workshop webinars in integrating NAMs into their respective regulatory frameworks. Each agency's approach reflects its unique mandate and the specific challenges it faces in toxicity testing and chemical risk assessment. These summaries provide insights into the innovative strategies employed, the challenges encountered, and the strides made towards a more efficient, humane, and scientifically robust regulatory process.

2.2.1 U.S. Environmental Protection Agency

U.S. EPA conducts toxicity testing to evaluate the safety of chemicals and their potential environmental impact. EPA's portfolio includes pesticides, fertilizers, industrial chemicals, pollutants, and substances that affect air and water quality. Standard testing often involves using laboratory animals to determine dose-response relationships, which inform about exposures that could cause harm and the threshold for adverse effects. EPA must often rely on data generated by pesticide manufacturers or chemical sponsors. These tests often assess acute exposure (short-term), chronic exposure (long-term), and special cases (e.g., developmental toxicity, neurotoxicity, and carcinogenicity). The amount of data available varies by chemical type. For instance, pesticides often have rich databases, as a large number of studies are required for pesticide registration, while other types of chemicals may have less-comprehensive datasets available. Based on available data, EPA establishes regulatory limits for chemical exposures, such as permissible levels in air, water, soil, and on agricultural produce.

In its strategic endeavor to modernize toxicological methods, EPA unveiled its NAMs Work Plan[8], first released in 2020 and subsequently refreshed in 2021, outlining a forward-looking roadmap underscored by five parallel, mutually reinforcing objectives. Central to the plan is the commitment to reducing animal testing while filling critical information gaps. First, the plan underscores the necessity for evaluating current regulatory flexibility to understand statutory and policy frameworks governing current toxicity methods and thereby discern the policy implications for integration of NAMs. Second, the work plan emphasizes establishing quantifiable baselines and metrics indispensable for monitoring advancements, particularly concerning the reduction in animal usage. The third objective revolves around solidifying scientific confidence in NAMs, evidenced by the agency's

6 https://www.oecd.org/chemicalsafety/risk-assessment/considerations-for-assessing-the-risks-of-combined-exposure-to-multiple-chemicals.pdf

7 See also the FDA Modernization Act https://www.congress.gov/bill/117th-congress/senate-bill/5002/text and a related editorial https://www.science.org/content/resource/new-path-new-drugs-finding-alternatives-to-animal-testing

8 https://www.epa.gov/system/files/documents/2021-11/nams-work-plan_11_15_21_508-tagged.pdf

ongoing efforts to construct a robust scientific confidence framework to assess the quality, reliability, and relevance of NAMs in regulatory decision making. The fourth objective focuses on the proactive development of NAMs tailored to address pivotal information gaps, necessitating discernment of which methods to prioritize and the decision-making contexts they inform. The final goal involves comprehensive stakeholder engagement, acknowledging that a transition to NAMs requires a harmonized, multi-sectoral effort spanning governmental departments, academia, industry, and international partners, fostering collaboration and optimizing resource utilization across diverse regulatory and research spheres.

EPA is actively promoting the integration and acceptance of NAMs through educational initiatives, stakeholder engagement, biennial conferences, and targeted training programs. As the agency shifts toward innovative techniques such as leveraging genomic, transcriptomic, metabolomic, and proteomic data, one goal is to enhance the scientific rigor and applicability of toxicity testing, particularly for data-poor chemicals, potentially shortening assessment timelines. Simultaneously, the agency is transforming its validation and reporting standards to be more NAM-compatible. This evolution is evident in international collaborations aimed at updating guidance documents and developing new reporting frameworks. EPA's recent proposals for alternative endocrine disruptor screening models[9] and the availability of advanced *in silico* tools[10] further demonstrate its commitment to making environmental protection strategies more efficient, accessible, and scientifically sound.

2.2.2 U.S. Food and Drug Administration

U.S. FDA's toxicity testing focuses on substances humans ingest, such as food additives and chemicals from packaging materials, drugs, and chemicals applied on the body, like cosmetics and personal care products. Like EPA, FDA uses sponsor data based on animal studies, *in silico* studies, and *in vitro* studies to assess the safety of these substances, looking for evidence of harmful effects including cancer, birth defects, and organ damage. The level of reliance on animal data varies depending upon the FDA center. For pharmaceuticals, data from non-clinical studies (e.g., *in silico*, *in vitro*, and *in vivo*) that support the safety of a drug are submitted to FDA before human clinical trials can begin and at later stages throughout drug development.

FDA has a long-standing commitment to promoting the development and use of NAMs to evaluate and predict the safety, effectiveness, and reliable manufacture of regulated products[11], seeking to address the limitations of animal testing and better model human toxicity. The Food and Drug Omnibus Reform Act of 2022 (FDORA) defined "*nonclinical test as a test conducted* in vitro*,* in silico*,* in chemico*, or a nonhuman* in vivo *test that can serve as supporting data for regulatory decision making*[12]*, and several guidance documents are also available concerning acceptable alternative methods*."[13] FDA's fiscal year 2023 budget included $5M in new funding to support the New Alternative Methods Program through FDA core operations[14]. The NIH Small Business Innovation Research Program also provides non-dilutive funding opportunities for "*organotypic models using cells from rat or mouse models or other experimental animal models, with a focus on comparisons between* in vivo *and* in vitro *toxicity endpoints*."[15] Funding opportunities are offered to early-stage research and development three times a year.

Additionally, the Center for Drug Evaluation and Research has published papers describing gaps and challenges around current nonclinical testing strategies (Avila et al., 2020, 2023). Furthermore, the Center for Devices and Radiological Health has established the Medical Device Development Tools Program to qualify tools for use in the development and evaluation of medical devices, including nonclinical assessment models[16].

Through strategic collaborative engagements, FDA is at the forefront of multi-stakeholder partnerships, both domestically and on the global stage. Activities include collaborations between the various FDA centers and industry/academia to develop airway, gut, cardiac, brain, testes, placenta, and liver microphysiological systems, human blood vessel chips, a neuromuscular junction chip, and the human intestinal microbiome, via research collaborative agreements and collaborative research and development agreements. Wang et al. (2021) reviews the application of 3-dimensional cell culture models for regulatory submissions. FDA places a strong emphasis on public communication and transparency – evident in its efforts to engage with stakeholders through clear, open channels that not only disseminate information but also invite public input. Such transparency aims to foster public trust and facilitate an informed dialogue on the implementation of NAMs, ensuring that the transition is both inclusive and responsive to societal expectations.

2.2.3 European Food Safety Authority

EFSA, an agency of the European Union established in 2002, oversees similar food- and feed-relevant substances as FDA, but operates under regulatory standards pertinent to EU member countries. Toxicity testing is used to evaluate chemicals related to food and

9 https://www.regulations.gov/document/EPA-HQ-OPP-2021-0756-0002

10 https://www.epa.gov/comptox-tools/toxicity-estimation-software-tool-test

11 See, for example, https://www.fda.gov/media/144891/download?attachment

12 https://www.thefdalawblog.com/wp-content/uploads/2023/01/HPM-FDORA-Summary-and-Analysis.pdf

13 See for example: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/assessing-credibility-computational-modeling-and-simulation-medical-device-submissions; https://www.fda.gov/regulatory-information/search-fda-guidance-documents/s5r3-detection-reproductive-and-developmental-toxicity-human-pharmaceuticals; and https://www.fda.gov/regulatory-information/search-fda-guidance-documents/oncology-pharmaceuticals-reproductive-toxicity-testing-and-labeling-recommendations-guidance

14 https://www.fda.gov/science-research/about-science-research-fda/advancing-alternative-methods-fda

15 https://seed.nih.gov

16 https://www.fda.gov/medical-devices/medical-device-development-tools-mddt

feed safety (e.g., food additives, flavorings, packaging materials, and other food and feed improvement agents and contaminants), plant-protection products (e.g., pesticides), novel foods, nutrient sources, and genetically modified organisms. The processes also rely heavily on animal testing to determine safe levels of exposure for humans. EFSA conducts risk assessments based on these tests to inform regulatory limits (acceptable daily intakes or tolerable upper intake levels) for various chemicals and substances.

EFSA is actively navigating the rapidly evolving landscape of regulatory toxicology, particularly in addressing new challenges posed by novel foods, nanomaterials, and chemical mixtures in the environment (EFSA Scientific Committee, 2021a,b; EFSA NDA Panel et al., 2021). These complexities, coupled with a surge in public consciousness about chemical exposure risks, have catalyzed EFSA's strategic shift towards integrating NAMs into its risk assessment paradigm. Influenced by broader policy initiatives like the European Green Deal and the EU Chemicals Strategy for Sustainability, EFSA's 2027 strategy[17] explicitly advocates for integrating NAMs to reduce animal testing, as well as to enhance the quality and efficiency of chemical risk assessments. EFSA acknowledges an urgent need for validated NAMs capable of accurately predicting chemical absorption and bioavailability, thereby reducing unnecessary *in vivo* testing while safeguarding public health. This evolution aligns with the organization's pronounced commitment to collaboration both at the European and international levels.

EFSA has already invested considerably in moving toward the use of NAMs in its regulatory processes, directing around €20 million into projects and case studies that support the integration of NAMs into guidance documents like that on the use of read-across for food safety assessment[18], focusing on critical research gaps that NAMs could fill.[19] Specifically in terms of understanding absorption, toxicokinetics, or *in vivo* kinetics, EFSA guidance recommends a tiered approach applicable to substances including food additives and nanoparticles, in which absorption is a decision point for higher-tiered studies.

2.2.4 Spanish National Environmental Health Centre

The Spanish National Environmental Health Centre (Centro Nacional de Sanidad Ambiental; CNSA) is part of the Instituto de Salud Carlos III (Carlos III Health Institute), a research institution that provides scientific and technical services and educational programs supporting the national health system. CNSA findings play a crucial role in developing national guidelines and policies for environmental health and safety. CNSA's focus extends to chemical substances present in the environment and consumer products, including pesticides, heavy metals, and endocrine-disrupting chemicals. Through the Risk Assessment Unit, CNSA is developing new conceptual approaches for assessing the impact of chemical exposure on human health, combining information from traditional toxicity testing methods, NAMs, and human biomonitoring for assessing human risks posed by potentially hazardous substances.

CNSA's participation in the EU's expansive Partnership for the Assessment of Risks from Chemicals (PARC)[20] highlights a concerted move towards advanced risk assessment strategies, with Work Packages 5 (Hazard Assessment) and 6 (Innovation in Regulatory Risk Assessment) being particularly relevant (De Castelbajac et al., 2023; Marx-Stoelting et al., 2023). Although current efforts under PARC notably lack a specific focus on the gut or metabolism, many of the principles are readily applicable.

In the area of chemicals present in foods, pesticide residues and nanoplastics are a priority for risk assessment. NAMs can provide valuable mechanistic information regarding hazards such as neurotoxicity (Martin-Folgar et al., 2024), which are essential for accurate risk assessment but have limited coverage by animal studies. The integration of dietary exposure models and biomonitoring offers options for improving risk assessment methods and addressing individual variability (Tarazona et al., 2022). CNSA is also reviewing recent developments, including evidence-based methods (Hoffmann et al., 2022), and developing a new conceptual approach for the assessment of adversity and individual variability in refined adverse outcome pathway networks.

Looking beyond methodologies, CNSA advocates for a revolutionary shift in risk assessment. This entails not just incorporating alternative methods into the current system, but developing entirely new paradigms for safety assessment based on integrated approaches and the mechanistic understanding that can be provided by NAMs. The final aim is to achieve risk characterization approaches addressing human variability and covering vulnerable population groups for informing public health strategies.

2.3 Summary/strategic directions

Regulatory agencies are experiencing an era of transformation in risk assessment as they move toward a holistic paradigm that neither strictly adheres to animal models nor completely depends on NAMs. Instead, in an ideal scenario, the strengths of both will be leveraged, possibly in a tiered approach, to facilitate an evolution in risk assessment away from traditional *in vivo* studies. A shift towards NAMs requires a careful balancing act for regulatory agencies – traditional methods are embedded within established legal frameworks and regulations, and the transition to NAMs necessitates a detailed assessment of this landscape and a push to update existing guidance documents. Integrated approaches to testing and assessment (IATA) can be employed to maximize the use of available data and the use of NAMs for addressing the remaining information gaps (Cattaneo et al., 2023).

Concurrently, there is clear intent to harmonize standards with other bodies to create a unified regulatory environment within the United States and beyond. The pursuit of cross-agency, cross-sec-

17 https://www.efsa.europa.eu/sites/default/files/2021-07/efsa-strategy-2027.pdf

18 https://www.efsa.europa.eu/sites/default/files/wgs/cross-cutting-science/wg-read-across.pdf

19 https://www.eu-parc.eu/

20 https://endallergiestogether.com/research/food-allergy-statistics/

tor collaboration and global partnerships facilitates consensus on accepted methods, data standards, and protocols, which can aid worldwide acceptance of NAMs – both in regulatory contexts and for other uses. Fostering broad engagement of both stakeholders and the public through transparent dialogue and education can also elevate the credibility of novel methods.

Looking ahead, agencies recognize that advanced technologies like artificial intelligence (AI), machine learning, and quantum computing will be pivotal for enhancing validation processes and facilitating comprehensive and nuanced risk assessments – work that is particularly important for data-poor chemicals. The overarching goal is to ensure that, as science progresses, the regulatory environment adapts accordingly, with a continued focus on human and environmental safety, ethical considerations, and scientific integrity.

3 Evaluating potential allergenicity

Food allergies are a significant and growing public health concern, affecting millions of people worldwide. Food allergens represent a unique class of substances that interact intimately with the GIT. The ability of these allergens to cross the intestinal epithelium and induce immune responses places them at a critical juncture of gastrointestinal and systemic health.

The mechanism of food allergy is complex. In IgE-mediated allergies, the immune system identifies an otherwise innocuous food protein with allergen-specific IgE antibodies. Upon subsequent exposures, these allergen-specific IgE antibodies cross-link to the allergen and to immune effector cells (e.g., mast cells, basophils), leading to their activation/degranulation and release of allergic mediators. Clinical symptoms can range from oral and dermal itching, nausea, and abdominal pain, to life-threatening anaphylaxis.[21] Other types of food allergies, including eosinophilic gastrointestinal diseases, are mediated through different adaptive immune mechanisms through which allergens can be recognized by allergen-specific T cell subsets that secrete allergic cytokines, resulting in the recruitment of eosinophils to gastrointestinal sites (Licari et al., 2020). The precise mechanisms by which food antigens cross the intestinal barrier and the ensuing immune responses are active areas of research.

The accurate assessment of food allergens is paramount in regulatory contexts, and animal models have limitations in accurately replicating human immune responses to food allergens (Kazemi et al., 2023). Integration of NAMs could improve the evaluation of allergenic potential and enhance understanding of the systemic implications of food allergens, which could enhance regulatory decisions, safety assessments, and public health strategies. NAMs may provide a more nuanced and human-relevant framework for understanding the complex interplay between diet, the GIT, and systemic health, which is essential for addressing the challenges posed by food allergies in both clinical and regulatory settings.

Two pre-workshop webinars provided relevant background information on food antigen transport across the intestinal epithelium and structural features of allergen-specific antibodies – topics highlighting the need for a more comprehensive understanding of food allergens and their impact on human health.

3.1 Food antigen transport across the intestinal epithelium: Outcomes and consequences

Simon P. Hogan, Askwith Research Professor of Food Allergy at the University of Michigan, explained the current knowledge of how food antigens cross the intestinal epithelium and the differences in antigen movement between healthy individuals and those with food allergies.

The small intestine, with its highly specialized epithelial layer designed for nutrient absorption and defense, is the primary focal point for the degradation of complex nutrients into absorbable forms – a process facilitated by the extensive surface area provided by villi and microvilli. The intestinal epithelium comprises various cell types, such as Paneth cells for microbial defense, goblet cells for mucus secretion, enteroendocrine cells for hormonal regulation, and enterocytes for lipid, protein, and carbohydrate metabolism and absorption. The healthy epithelium acts as a selective barrier, permitting beneficial substances to enter the body from the lumen while blocking toxins and pathogens. Disruption of this barrier is associated with both GIT and systemic autoimmune diseases, such as inflammatory bowel disease, metabolic disorders, and food allergies. The content of the gut lumen also plays a pivotal role in immune education and the establishment of oral tolerance to dietary antigens.

To investigate the process by which dietary antigens cross the intestinal epithelium, Hogan's team employed two-photon imaging on live, anesthetized mice to track the movement of fluorescently labeled food antigens. Immunostaining indicated that goblet cells rapidly and indiscriminately take up labeled antigens. Goblet cell antigen passages (GAPs) are known to pass luminal antigens to immune cells, and GAP expression is spatially and temporally regulated within the intestine (McDole et al., 2012; Kulkarni et al., 2020; Newberry and Hogan, 2021; Gustafsson et al., 2021; Noah et al., 2021).

Under food-allergic conditions, antigen passage patterning and landscape is dysregulated, and multiple intestinal epithelial cell populations of the secretory lineage contribute to antigen uptake. These findings indicate that food allergens translocate across the small intestinal epithelium via non-canonical antigen passages, which Hogan's group termed secretory antigen passages (SAPs). Experiments showed that, in allergic mice, antigens were presented to mast cells within minutes, implicating SAPs as critical in the rapid allergic response process. Crucially, genetically engineered mice that were unable to activate SAPs did not exhibit IgE-mediated allergic responses as measured by systemic shock, highlighting the potential for new treatment approaches that target these passages.

[21] https://www.hopkinsmedicine.org/health/conditions-and-diseases/food-allergies

Further exploration into the immune environment revealed that alterations in IL-4 and IL-13 signaling pathways, known to be linked with food allergies, could convert GAPs into SAPs. This suggests a genetic predisposition to food allergies, in which mutations in IL-4Rα, which lead to heightened signaling, result in an increased number of non-goblet cells acting as antigen passages. This genetic link was strengthened, as mice engineered with the IL-4Rα mutations demonstrated food sensitization and subsequent allergic reactions upon allergen exposure.

The research was extended to human models using human intestinal organoids developed from pluripotent stem cells and transplanted into mice. These organoids formed normal human intestinal tissue, including goblet cells capable of antigen passage. When IL-4R was activated in these organoids, both an increase in GAPs and contributions from other secretory cells were observed, mirroring the mouse model findings and reinforcing the idea that a variety of intestinal cells can participate in antigen passage and that these mechanisms may be key to understanding and potentially controlling food allergies.

3.2 The induction of allergen-specific antibodies in allergy and oral tolerance

Sarita Patil, assistant professor at Harvard Medical School, described her group's work related to the structural features of allergen-specific antibodies.

When patients with food allergies are exposed to even minute amounts of causative food antigens, allergen-specific IgE cross-linking is triggered on the surface of mast cells, stimulating degranulation and initiating a multi-systemic response that can include anaphylaxis. Allergic individuals likely have pathogenic T cells that present allergens in the presence of pro-Th2 cytokines like IL-4 and IL-13, which drive antibody class switching of memory IgG1 cells to IgE, predisposing individuals to anaphylaxis.

While oral immunotherapy, through careful increases in allergen exposure, can shift the immune system from producing IgE to IgG, allergen-specific IgG levels alone do not explain tolerance achieved through immunotherapy. In clinical trials, similar elevations in serum allergen-specific IgG4 levels are seen in peanut-allergic individuals with both sustained and transient tolerance to peanuts. Thus, focusing on the peanut allergen Ara h 2, Patil's group considered qualitative differences in antibodies, like affinity or epitope recognition, that could account for the sustained tolerance seen in some patients.

Using techniques ranging from sequential epitope mapping to X-ray crystallography, they discovered several antibodies found exclusively in patients with sustained tolerogenic responses, suggesting that certain antibody-epitope interactions might be key to long-lasting tolerance. Notably, mAbs that targeted conformational epitopes were more effective at inhibiting polyclonal IgE recognition of Ara h 2 than were antibodies that recognize sequential epitopes. These data indicate that patients with sustained clinical efficacy have unique high-affinity antibodies that can block multiple IgE epitopes, both sequential and conformational, inhibiting IgE-mediated activation of effector cells and providing sustained clinical efficacy post immunotherapy.

The researchers also uncovered an intriguing pattern of "convergent evolution" in immune responses to the Ara h 2 allergen. They observed strikingly similar B-cell receptor sequences across patients, indicating a stereotypical, highly homologous immune response to Ara h 2, which was subsequently confirmed by other groups (Patil et al., 2015; Croote et al., 2018, 2024; Hoh et al., 2020). Patil's group showed that these convergent antibodies are not just convergent by sequence, but also in the way they recognize allergens, indicating that, on a population basis, many people recognize the immunodominant Ara h 2 in a highly similar fashion.

Further investigations involved using an IgE-inhibition ELISA to understand the mechanism of neutralization within a larger group of peanut-allergic patients treated with oral immunotherapy. The study identified neutralizing IgG4 antibodies in patients with sustained responses, which blocked multiple IgE epitopes and sequestered the allergen away from effector cells, preventing IgE-mediated activation. *In vivo* experiments with a humanized mouse model confirmed that these neutralizing antibodies provided more protection against allergic responses than non-neutralizing antibodies.

To reduce the allergenicity of Ara h 2, a hexamutant version with mutations in identified conformational epitopes was engineered. This mutant demonstrated impaired binding to serum IgE in patient serum *in vitro* and provoked a lower allergic response in the murine anaphylaxis model, pointing towards a potential path for creating less allergenic versions of Ara h 2 (Min et al., 2024).

4 Gastrointestinal models: State of the science

During the in-person workshop, speakers from academia, industry, and regulatory bodies explored the state of the science regarding the use of GIT NAMs to examine questions of absorption, metabolism, and toxicity. This section summarizes those presentations and provides insight into the readiness of these innovative tools to fill critical research gaps, with the aim to either complement or replace current risk assessment methods. These talks were meant to showcase the possibilities of various model systems, highlighting their strengths and limitations, which may help to establish the scientific confidence needed for their incorporation into regulatory frameworks.

In vitro toxicological models range from simple cell line-based monolayers, which primarily assess functional aspects such as absorption and toxicokinetics, to 3D co-cultures with improved physiological relevance, which can potentially address mechanistic questions. Choosing a GIT model that is fit for purpose is critical, as there is often a tradeoff between physiological relevance, scalability, cost, and ease/speed of implementation. The selection process must weigh criteria including the specifics of the biological samples needed to answer the research question, the need for specific cellular interactions, and the need to incorporate mechanical properties such as flow or stretch.

For the purpose of this report, talks are arranged based on the type of *in vitro* model employed: "simple" cell culture models including monolayers and transwell setups, gut-on-a-chip models

that allow the incorporation of aspects such as flow and stretch, and 3D intestinal organoids, which reproduce some important aspects of the cellular architecture and functional complexity of the human intestine.

4.1 Monolayer and transwell cell culture models

Traditional *in vitro* cell culture models of the human GIT, encompassing both monolayer and transwell setups, are instrumental in pharmacokinetic and toxicokinetic studies, providing insight into drug absorption, gastrointestinal toxicity, barrier function, and the impact of substances on cellular metabolism, signaling pathways, and gene expression. These models can include either immortalized cell lines or human primary cells. Caco-2, an immortalized cell line derived from human colon carcinoma, is one of the most-used cell lines in GIT research (Fogh et al., 1977). Caco-2 cells can spontaneously differentiate and form monolayers that closely mimic small intestinal enterocytes in the intestinal mucosa, making them invaluable for studying intestinal absorption and transport processes. Despite the advantages of ease and standardization provided by a cell line-based model, Caco-2 cells alone do not represent the cellular diversity of the human intestine, and they may exhibit altered expression of some transporters and enzymes compared to normal enterocytes (Ölander et al., 2016). Primary human cells, typically sourced directly from human tissue and generally inclusive of various cell types, offer a more physiologically relevant model. However, tradeoffs include limited availability and potential variability between samples.

While some GIT cell culture models use simple monolayers, other studies may employ transwells, which include a physical barrier between two compartments, mimicking the separation between the intestinal lumen and the submucosal layers or vascular space. Transwell studies are particularly useful for studying transport across epithelial barriers, drug absorption, and GIT toxicity, allowing for the interaction between cell types.

4.1.1 Using Caco-2 permeability to estimate oral bioavailability for environmental chemicals

In the context of environmental chemicals, oral bioavailability refers to the extent and rate at which an environmental pollutant is completely available to the systemic circulation of a living organism (Price and Patel, 2024), encompassing the processes of absorption through the GIT wall and fraction of dose escaping first-pass elimination in the GIT wall (enterocytes) and liver (Fedi et al., 2021). Compared to pharmaceuticals, studying the bioavailability of environmental chemicals is complicated due to the inability to test them on humans and poor interspecies correlation of quantitative *in vivo* oral bioavailability (Musther et al., 2014). Elaina Kenyon, a research toxicologist at EPA, delivered a presentation describing EPA's use of Caco-2 cells to increase the accuracy of bioavailability predictions.

Caco-2 cells were chosen due to their accessibility, established presence in pharmaceutical research, and their known ability to measure apparent permeability (P_{app}) in a way that correlates with effective permeability (P_{eff}) *in vivo* (Dahlgren et al., 2015; Yim et al., 2020). Measurement of P_{app} allows estimation of the fraction absorbed component of bioavailability (Wetmore et al., 2012; Punt et al., 2022). The study measured both apical-to-basolateral and basolateral-to-apical permeability for a panel of chemicals from the ToxCast chemical library (Richard et al., 2016) for which bioactivity data and *in vitro* toxicokinetic data were available (Wambaugh et al., 2019). Barrier integrity was monitored by transepithelial electrical resistance (TEER) and lucifer yellow permeation. Reference chemicals for low (ranitidine) and high (warfarin) permeability and active efflux (talinolol) were run on each transwell plate. Bidirectional P_{app} (P_{app}^{AB}, P_{app}^{BA}) was determined, and efflux ratio (ER) was calculated from these data (Honda et al., 2024).

For P_{app}^{AB} (used to estimate fraction absorbed), the assay achieved measurable P_{app} for 474 out of 484 chemicals, with 310 chemicals remaining after filtering based on fractional recovery (F_{rec}). P_{app}^{BA} measurements provided measurable P_{app} for 478 out of 484 chemicals, with 396 chemicals remaining after filtering based on F_{rec}. Further, ERs were calculated for 302 chemicals to determine which were likely subject to active transport. These data are useful for comparison against pharmaceutical data, building quantitative structure-property relationship (QSPR) models for fraction absorbed, and incorporation into existing frameworks (e.g., high-throughput toxicokinetics (HTTK), see 4.1.2) to provide refined estimates of oral bioavailability. Such information can improve screening-level risk prioritizations and reduce the uncertainty associated with IVIVE of bioactivity screening data using *in vitro* toxicokinetic methods (Honda et al., 2024).

Kenyon emphasized that overall oral bioavailability is a multicomponent process that includes more than just the fraction absorbed through the gut wall – the focus of many *in vitro* assays. Bioavailability also includes the fraction that escapes first-pass elimination in both the gut and the liver, and it is influenced by gastrointestinal physiology and physicochemical properties of the chemical itself, as well as numerous host factors such as diet, life stage, and pre-existing health conditions.

She concluded by showcasing an application of their findings in a health risk prioritization context, explaining that incorporating bioavailability data can significantly refine bioactivity predictions and reduce uncertainties associated with translating *in vitro* data to real-world scenarios. Kenyon noted that these Caco-2-derived data will soon be publicly available to contribute to broader scientific and risk assessment endeavors.

4.1.2 Modeling oral bioavailability of environmental chemicals

John Wambaugh, a research physical scientist from EPA, discussed the integration of *in vitro* gut data into a risk prioritization model for environmental chemicals. Three components are needed to consider the public health risk of a chemical: the hazard, the exposure, and the dose-response relationship (NRC, 1983). While biological hazard has traditionally been assessed using animal studies, *in vitro* high-throughput screening can be used as an alternative measure of *in vitro* bioactivity. To translate *in vitro* data to meaningful *in vivo* predictions, however, *in vitro* bioactivity concentrations must be linked to potential *in vivo* toxic doses

via IVIVE (Rotroff et al., 2010; Coecke et al., 2013; Wetmore et al., 2015). However, IVIVE requires toxicokinetic data, which are notably lacking for the vast majority of non-pharmaceutical chemicals (Bell et al., 2018).

To address this gap, HTTK is being developed, adapting pharmaceutical industry techniques to estimate toxicokinetics for environmental chemicals. These methods aim to provide ballpark toxicokinetic data for large numbers of substances, which could subsequently be used for IVIVE. HTTK combines rapid *in vitro* measurements (possible for hundreds to thousands of chemicals) with generic toxicokinetic models capable of utilizing these data for predictions. The end goal of the complex process is to calculate what is termed the "administered equivalent dose" – essentially, the specific dose required to achieve a certain *in vitro* concentration within a living organism – linking the *in vitro* findings to potential real-world exposure levels. Wambaugh emphasized that HTTK essentially identifies a single scaling factor, making it possible to predict how much of a chemical is needed to be administered to reach a specific concentration in the blood, considering the most sensitive individuals in the population. These data allow for rapid chemical prioritization, based on the likelihood that a portion of the population is exposed to the bioactive dose.

Integration of the *in vitro* Caco-2 gut absorption data presented by Elaina Kenyon (see 4.1.1) into the chemical prioritization scheme expanded the data set significantly. However, gaps remain, prompting the team to develop a QSPR model using machine learning (specifically, a random forest method) to predict chemical permeability. While not perfect, this model offers reasonable predictions for chemicals without direct Caco-2 measurements, particularly distinguishing between high, medium, and low permeability (0.2×10^{-6} cm/second, 2×10^{-6} cm/second, and 20×10^{-6} cm/second, respectively).

QSPR data were incorporated into the HTTK model, which already includes factors like first-pass metabolism. The improved model now accounts for the fraction of a substance absorbed from the gut and escaping initial metabolism, enhancing predictive capability. Comparing the updated model against traditional *in vivo* rat studies and other tools like ADMet predictor[22], Wambaugh asserted that their approach holds up well, even surpassing the predictive power of rat studies in certain respects. For example, the root mean squared error human systemic bioavailability compiled by Kim et al. (2014) was 0.08 for predictions based on Caco-2 *in vitro* measurements, 0.08 for predictions based on Caco-2 QSPR predictions, and 0.26 for rat-based *in vivo* measurements (Musther et al., 2014; Wambaugh et al., 2018).

Wambaugh emphasized that the enhanced HTTK model can allow risk assessors to better determine the margin between actual exposure and potentially bioactive doses. Although the model has noted limitations and uncertainties, it provides valuable insight, especially when direct chemical data are lacking, and continues to refine the risk assessment process for environmental chemicals.

4.1.3 Characterization of a human *in vitro* intestinal model for the hazard assessment of engineered materials

In her presentation, Christie Sayes, a professor from Baylor University, described work on creating a transwell co-culture model of the human gut that simulates realistic interactions within the intestinal environment and could provide a reliable, high-throughput alternative to animal testing. The model includes undifferentiated and differentiated Caco-2 cells to represent the intestinal lining, HT-29 cells to represent goblet cells, and Raji B cells to mimic the immune compartment. This *in vitro* model involves a 21-day procedure with a two-phased approach: the initial adaptation of Raji B cells into media representative of the entire triculture model, and the subsequent differentiation of Caco-2 cells, induced by Raji B cells, into M cells, concluding with the addition of mucus-secreting HT29 cells.

Sayes and her team are working to optimize three key endpoints: barrier permeability, antigen transport, and inflammatory responses (Gibb et al., 2021). They utilize various microscopy techniques (e.g., bright field, scanning electron, and confocal) to quantify cellular characteristics and responses and compare those data to a particular exposure scenario. Detailed morphological characterization aids in understanding the specific impacts of various toxicants, providing a comprehensive view of cellular behavior postexposure. This careful imaging, along with functional analysis (e.g., expression of cytokines and cell-surface receptors; TEER), established the optimal passage number (20 to 30) and the optimal day to expose the system to a toxicant (day 8 to 11).

In vivo, substances that enter through the GIT undergo various phases of digestion prior to absorption in the small intestine. To simulate these conditions, samples underwent a three-phase simulated digestion before being added to the model system, including an oral phase, gastric phase, and intestinal phase (Ede et al., 2020; Pradhan et al., 2020). During each phase, pH, enzymes, and salts were included over a 4-hour period to mimic *in vivo* conditions (Minekus et al., 2014; DeLoid et al., 2017). Enzymatic gastric digestion changed the surface texture of particles such as microplastics and cellulose fibers, increasing surface roughness, which could influence cellular inflammatory responses (Sayes et al., 2007; Pradhan et al., 2020). After undergoing simulated gastric fluid digestion, cellulose nanocrystals became slightly frayed; however, the crystals induced similar pro-inflammatory responses as undigested cellulose nanocrystals or conventional cellulose.

Sayes underscored the necessity for high-throughput *in vitro* methods within toxicological studies, advocating for a mixture-based approach, which is critical for navigating the complexities of real-world exposure scenarios such as the combination of toxicants that might be present in drinking water.

4.1.4 Evaluating the technical quality of a triculture gut model to test particle permeability

Validating NAMs for commercial or regulatory use is a complex, multicomponent process, one aspect of which includes detailed technical characterization.[23] Elijah Petersen, a research scientist

[22] https://www.simulations-plus.com/software/admetpredictor/

[23] https://ntp.niehs.nih.gov/sites/default/files/2023-09/VWG_Doc_Comments_ACC_20230905.pdf

from the National Institute of Standards and Technology (NIST), delved into the intricacies of the technical characterization process, referring to a collaboratively produced framework providing guidelines for enhancing the technical quality of NAMs in preparation for potential commercial or regulatory applications (Petersen et al., 2023). The framework is non-linear and potentially iterative, aiming to generate NAMs that provide reproducible results across time and among laboratories.

Following this framework, Petersen and his team are analyzing multiple aspects of a tri-culture gut model designed to test particle permeability. Parameters being assessed to understand and optimize the model include cell viability and metabolic activity (MTS assay), mucous production (microscopy/ELISA), barrier integrity (TEER), and permeability of reference compounds.

Extensive studies of barrier integrity illustrate that TEER measurements are particularly challenging. In the hands of the Petersen group, TEER measurements varied between operators and also exhibited day-to-day inconsistencies. Petersen noted that some of the variation researchers experience in TEER measurements could result from the devices themselves, including a lack of standardized procedures and guidelines. His group performed a detailed, anonymized comparison among different prong-based and chamber-based test systems. The work revealed different sources of variability, such as operator-to-operator and experiment-to-experiment variability, through evaluating KCl calibration data and data from the tri-culture model. Further work is being conducted to investigate approaches to reduce variability and maximize confidence in longitudinal data.

Efforts are ongoing to standardize TEER measurement practices to quantify certain assessments like mucous evaluation (which, currently, is only qualitative) and to address numerous other challenges inherent to *in vitro* models. Petersen emphasized the need for improved methods and understanding in these areas, noting that thorough technical characterization can provide the underlying measurement confidence necessary for broader application of this model.

4.1.5 A three-layer intestinal model for toxin translocation studies

Angela Melton-Celsa and Kristin H. Gilchrist, researchers from the Uniformed Services University's Department of Microbiology and Immunology and Center for Biotechnology, respectively, presented their collaborative work developing a three-layer intestinal model to study the translocation of Shiga toxins (Bova et al., 2023). The research holds particular relevance for the Department of Defense because Shiga toxin-producing *E. coli* can cause diarrheal and intestinal disease that can negatively impact troop readiness and pose serious health risks, including life-threatening hemolytic uremic syndrome. To develop effective treatment strategies, there is a need for a more nuanced understanding of how Shiga toxins (Stx) 1 and 2 cross the intestinal epithelial barrier.

Gilchrist described that, while conventional *in vitro* models such as Caco-2 monolayers indicate that transcellular transport is the most common mechanism for Stx translocation, evidence suggests that the behavior of Stx changes in more physiologically complex settings. Thus, to mimic the human intestinal environment more accurately, an innovative three-layer primary cell model was established in a transwell system. The transwell was separated by a PET membrane with a collagen/gelatin extracellular matrix on the apical side and a gelatin coating on the basal side. Myofibroblasts, differentiated from human adipose-derived mesenchymal stem cells, were grown in the apical matrix, and a primary colonic epithelial layer was seeded on top. The basal side contained a layer of human primary colonic microvascular endothelial cells. The model was carefully validated for cell viability, morphology, and expression of appropriate cellular markers, demonstrating successful co-culture conditions and promising preliminary results.

Melton-Celsa continued the discussion, focusing on experimental findings comparing the translocation of Stx in a single-cell-layer model containing only primary colonic epithelial cells versus the three-layer model. Initial tests on the single-cell-layer model showed similar translocation of Stx1 and Stx2 across the cell layer, irrespective of whether the toxins were placed on the apical or basolateral side. In the three-layer model, while both toxins still translocated the membrane in both directions, there was an overall 10-fold increase in toxin translocation compared to the single-layer model, which underscores the necessity for more complex models to understand translocation of these toxins in the gut.

Infection models were attempted with two strains of Stx-producing *E. coli*, and although strain 0157:H7 destroyed the model, O26:H11 did not, and low levels of toxin translocation were observed.

Future studies will aim to enhance understanding of the behavior of these toxins in complex environments, with a focus on improved understanding of the three-layer model and exploring how gut conditions, like the presence of elastase, could potentially modify the toxins, increasing their affinity for gut receptors – a factor that current models do not consider.

4.1.6 The development of a 96-well plate-based model of the human intestinal epithelium with applications for modeling toxicity and gastrointestinal pharmacokinetics

In his presentation, Bill Thelin, chief scientific officer at Altis Biosystems, described the company's work developing next-generation human intestinal primary cell models. Ultimately, Altis is working towards standardization of human primary cell cultures, integrating a genetically diverse pool of donor tissue, sourcing stem cells from multiple regions of the small and large intestines of healthy donors. Cells are expanded to ultimately produce commercial-scale lots, and quality controlled to ensure reproducibility. These primary cells can then be used across diverse platforms, including the RepliGut® system, a transwell-based model that simulates the entire life cycle of the native GIT, from proliferating stem cells through differentiation into populations of enterocytes, goblet cells, and enteroendocrine cells. The model is designed to allow researchers to thaw cells and move straight into experiments with minimal preparation.

Recognizing that genetic and phenotypic drift can occur with cell expansion, extensive studies were conducted to establish ideal passage numbers for each gut region. Altis employs stringent quality control measures to ensure both the phenotypic stability

and performance (measured using TEER profiles) of each cell lot produced.

Further, Altis is developing an efficient, high-throughput toxicology assay to predict GIT toxicity, which is often missed in non-clinical studies. This process involves a 96-well plate-based system that evaluates drug effects on cell cultures within 5 days, examining factors like cell viability, barrier formation, and cell proliferation. Initial studies with idarubicin (a DNA intercalator) and bortezomib (a proteosome inhibitor) demonstrated that the assay can distinguish impacts on proliferating versus non-dividing cells, providing a nuanced understanding of the underlying mechanisms of drug toxicity.

In an attempt to establish formal drug metabolism and pharmacokinetic (DMPK) modeling, Altis is using their primary cell transwell system to model drug absorption by the small intestine. Transcriptomic data were leveraged to ensure the model's physiological relevance, highlighting expression profiles of relevant genes, including the enzymes CES1, CES2, and CYP3A4. Collaborating with Genentech, Altis evaluated permeability using known pharmaceuticals, noting a reliable correlation with human data. Ongoing studies include analyzing drug transport via various pathways.

Finally, Altis is working with CN Bio to couple their model with a human liver model, aiming to simulate human oral bioavailability more accurately by including first-pass metabolism. This complex system allows observation of a drug's journey, including metabolism and transport, providing a more holistic view of how substances are processed in the body. The system's potential was demonstrated using temocapril (an ACE inhibitor), underscoring the benefits of more physiologically relevant systems, including accurate representation of enzyme expression.

4.2 Gut-on-a-chip technologies

The advent of gut-on-a-chip technologies marked a significant advance in gastrointestinal research, offering a more dynamic and physiologically relevant alternative to less nuanced *in vitro* methods such as Caco-2 monolayers and transwell studies (Marrero et al., 2021; Valiei et al., 2023). While Caco-2 cells, extensively utilized to predict human intestinal drug absorption, undergo spontaneous differentiation and form a tight intestinal monolayer, their inherent limitations in replicating the complex physiology of the human intestine have driven the development of advanced, microphysiological systems (MPS) (Bein et al., 2018; Costa and Ahluwalia, 2019; Lopez-Escalera and Wellejus, 2022). This advance is pivotal as regulatory agencies and industry partners seek reliable, accurate models that mimic human physiology more closely than past models. In preclinical investigations, it is important for pharmaceutical companies to evaluate the intestinal permeability of newly developed chemical compounds (Falcón-Cano et al., 2022). This is because novel drugs with unfavorable ADME profiles face an elevated risk of failure in clinical trials (Fedi et al., 2021). Unlike *in vitro* approaches that primarily rely on static cultures of cells in two-dimensional environments, MPS such as gut-on-a-chip models employ microfluidic technology to recreate the complex, three-dimensional architecture of the human GIT, serving as *in vitro* alternatives to animal models for testing the efficacy and safety of drug candidates with potential for improved predictivity. This not only allows for better simulation of the gut environment, including aspects like tissue stretch and luminal flow, but also facilitates the co-culture of multiple cell types (Kang and Kim, 2016; Bein et al., 2018; Thompson et al., 2020; Kopec et al., 2021). Thus, gut-on-a-chip models can more accurately mimic the intricate cellular interactions, barrier functions, and absorption processes of the human GIT, thereby refining predictive capabilities and providing a more representative model for studying the pharmacological and toxicological properties of ingested substances.

4.2.1 Evaluation of human intestinal epithelium in an MPS platform as an *in vitro* model for drug absorption

Ye Eun Jeong, an ORISE fellow at FDA, shared preliminary data from a project evaluating the regulatory potential of a gut MPS platform as a superior alternative to the static Caco-2 culture for drug absorption studies. Specifically, Caco-2 cells were grown under continuous fluidic conditions that enhanced nutrient and oxygen supply while facilitating waste removal[24] (Xiang et al., 2020). The group's primary question focused on whether microfluidic flow could induce Caco-2 cells to develop complex structures and biological barrier functions more closely resembling human gut physiology, thereby improving drug permeability predictions. This preliminary study examined several functional and morphological properties of the gut MPS model compared to a static Caco-2 monolayer. Additionally, gene expression profiles of Caco-2 cells, with an emphasis on drug-metabolizing enzymes and drug transporters, were explored.

Static Caco-2 monolayers are known to exhibit TEER values higher than those seen in the small intestine *in vivo*, implying a tighter and more intact *in vitro* barrier compared to the living intestine (Lopez-Escalera and Wellejus, 2022). Data from Jeong's group indicate that cells grown under dynamic flow have reduced TEER values and increased permeability of dextran compared to the static cultures, suggesting that microfluidic flow may help recapitulate physiologically relevant barrier integrity. Confocal immunofluorescence images show the expression of ZO-1, a tight junction marker (Lee et al., 2018), and F-actin, a key structural protein (Drenckhahn and Dermietzel, 1988), in both static and MPS groups. Cross-sectional views of Caco-2 cells constructed from z-stacked images suggest the formation of a thicker monolayer under a microfluidic environment. Furthermore, initial transcriptomic results support that microfluidic flow substantially upregulates the transcriptional levels of essential drug-metabolizing enzymes and drug transporters, including specific cytochrome P450 and solute carrier transporters, although the functional implications of this upregulation remain to be assessed.

Future comparative studies investigating transporter protein expression and their cellular functions in a gut MPS in comparison

[24] http://cn-bio.com

to the static model are needed to understand the potential of the microfluidic gut-on-a-chip model. Additionally, this examination could be extended to assess whether a gut MPS demonstrates an improved correlation with *in vivo* human drug absorption studies, encompassing drugs across all four Biopharmaceutics Classification System classes. These studies will provide insight into the regulatory potential of a gut MPS platform as an advanced *in vitro* model for drug discovery and regulatory sciences, surpassing the capabilities of traditional Caco-2 culture.

4.2.2 Biopsy-derived human intestine chips to investigate region-specific barrier responses

Ville Kujala, director of Discovery Biology at Emulate Inc., described Emulate's work to overcome the limitations of conventional two-dimensional cell models and organoids in mirroring human intestinal complexities using specialized duodenum and colon intestinal organ-on-a-chip technology. Emulate's system aims to provide a human-relevant approach to intestinal modeling, using tissue-specific biopsy-derived organoids and primary endothelium, and by incorporating tunable media flow and mechanical forces. The duodenum and colon chips contain two cell culture channels separated by a porous membrane. The top channel contains the intestinal epithelial cells, while the bottom channel contains the endothelial cells. Vacuum channels on the sides apply lateral stretch to the membrane.

Kujala presented confocal fluorescent imaging illustrating that major epithelial subtypes are present on the duodenum chip, including absorptive enterocytes, goblet cells, enteroendocrine cells, and Paneth cells, and that the chips exhibit proper epithelial polarization, correct location of major intestinal transporters, and tight barrier formation as assessed by apparent permeability measurements using dextran. Colon chips also exhibit major epithelial subtypes with expected donor-specific variability. In further characterizing the chip, experiments revealed that the presence of endothelium enhances the establishment of epithelial tight junctions and barrier formation. Additionally, epithelial polarity and cell height were improved by incorporating flow and stretch. Further, transcriptomic analysis using a novel metric called transcriptomic signature distance (Manatakis et al., 2021) illustrated that the transcriptome profile of colon intestine chip epithelium is significantly closer to *in vivo* colonic tissue than organoids in suspension culture. The endothelium was shown to play a key role in differentiating the transcriptomic signature.

In terms of application, Emulate has used colon intestine chips to model the dysregulated immune cell recruitment seen in inflammatory bowel disease by priming the chip with TNF-α and administering peripheral blood mononuclear cells (PBMCs) through the vascular channel. Robust and specific migration of PBMCs into the top channel was demonstrated compared to control, and the model was further shown to capture complex immune-mediated cytokine cascades and downstream barrier damage. Colon chips were used to assess efficacy of tofacitinib, a therapeutic used in inflammatory bowel disease. The chips captured the variable sensitivities of patient-derived organoids to the drug and illustrated that tofacitinib reduces proinflammatory signaling in response to IFN-γ. Further, early toxicity studies using the duodenum chip showed that treatment with a known GIT toxicant, indomethacin, resulted in a concentration-dependent increase in intestinal permeability and release of cellular injury markers, illustrating the chips' broad potential spectrum of applicability.

4.2.3 Accelerating synthetic biotic development with gut-on-a-chip technology

Tyler Nelson, from the 711th Human Performance Wing at the Air Force Research Laboratory, presented collaborative work describing a microphysiological system designed to generate predictive analyses and analyze countermeasures for optimizing the performance of pilots and other Air Force personnel.

A collaborative project with Synlogic aimed to assess engineered microbial strains (termed synbiotics) developed for potential treatment of phenylketonuria (PKU). The team used Emulate's Chip-S1® to create a microfluidic gut-on-a-chip model to test and evaluate the synbiotics under physiologically relevant conditions. In their model, the bottom "vascular lumen" compartment of the chip contained extracellular matrix proteins to allow for the attachment of vascular cells, while the upper "gut" channel contained Caco-2 and HT-29 cells. The model was matured under flow and cyclical stretch. Detailed microscopic analysis revealed relevant anatomy of both gut and blood compartments, including three-dimensional macrovillus-like structures that can provide surface area to retain microbes in the flow environment.

Synlogic's *E. coli* Nissle strain was engineered to detect phenylalanine (Phe) and convert it to a non-toxic biomarker, TCA. Using the model, the engineered strain was shown to lower Phe concentration, with a corresponding increase in biomarker production, in a dose- and time-dependent manner, and importantly, Phe levels in the vascular compartment were reduced (Nelson et al., 2021). A reduction in Phe was observed whether this amino acid was added to the gut compartment or the blood compartment, suggesting that the therapeutic strain applied to the gut alone could reduce circulating Phe levels. The study also demonstrated that supra-physiological doses of the bacterium had adverse effects on the model that were emulated at high doses in healthy volunteers. Further, extrapolations with data from non-human primates provided a correlation that supports the model's predictive capacity.

Further work involved utilizing synthetically engineered probiotics to potentially augment the gut-brain axis by eliciting a neurotransmitter, with the aim of increasing the resilience of warfighters in stressful environments. To do so, a replication-competent strain of *E. coli* Nissle was engineered to detect cortisol levels and, upon activation, produce an enzyme that metabolizes tryptophan (Trp) to tryptamine. Cortisol, Trp, and tryptamine were traced in both the gut and blood compartments, and the growth of *E. coli* Nissle was quantified as a function of time. This dynamic could be visualized using GFP tagging. Upon cortisol stimulation, Trp metabolism resulted in a spike in tryptamine production. Varying cortisol and Trp levels showed a linear relationship with the activation and efficacy of the genetic circuit. A considerable increase in cytokine activity was observed in response to Trp depletion, guiding the team to consider cytokine production as a feedback mechanism in the genetic circuits of future *E. coli* Nissle strains to prevent potential adverse effects (Nelson et al., 2023).

Future research aims to enhance the complexity of model systems by integrating intestinal enteroid-based cells and microfluidic devices to mimic a more physiologically relevant environment when a full complement of transporters and *in vivo*-relevant gene expression and metabolism are necessary. Incorporation of real-time oxygen sensors and a controlled atmospheric setting will allow the precise manipulation of the oxygen environment necessary for fecal microbiome studies, under varied conditions of diet, chemical/pathogen exposure, or extreme temperatures, further enhancing the understanding of host-microbiome interactions under physiological stressors.

4.3 Intestinal organoids

Intestinal organoids are three-dimensional, self-organizing structures derived from stem cells, capable of mimicking some of the cellular architecture and functional complexity of the human intestine. This inherent three-dimensionality allows for the study of intricate aspects of intestinal biology that are not readily replicated in the two-dimensional setups of transwells. Organoids are increasingly used for drug screening, efficacy testing, and toxicity risk assessment, often offering a more accurate prediction of drug responses and adverse effects in the human GIT. In conventional organoid culture, the apical (luminal) surface faces inward, and the basolateral surface faces outward. However, to access the apical surface, researchers can manipulate the organoid culture to flip the orientation, so the apical surface faces outward. This orientation flip is important for studies involving luminal materials such as nutrients, oral pharmaceuticals, and other ingested substances. The ability to create organoids in both apical- and basal-out orientations allows researchers to accurately model and investigate a wide range of intestinal functions and interactions.

4.3.1 Human stem cell-derived intestinal organoids as a tool to estimate human oral exposure and presystemic metabolism

In his presentation, Patrik Lundquist, a researcher from Uppsala University, discussed the work of his team, led by Per Artursson, which focused on establishing new *in vitro* methods to estimate the ADME properties of drugs and chemicals to assist with de-risking. The work involves using both traditional *in vitro* methods such as Caco-2 and new intestinal models to study permeability and presystemic metabolism, respectively. These assays are applied to biologics, small molecule drugs, and peptide drugs. Systems are categorized with global as well as targeted proteomics, and data are compared to clinical samples to aid in the creation of PBPK or toxicokinetic models.

To estimate intestinal permeability, the team examines directional and bidirectional transport in a transwell Caco-2 cell model, while also measuring drug concentration in the epithelium. One finding from this model indicates that substantial drug accumulation can occur within the cell layer, highlighting differences between cellular uptake and transepithelial permeability that have implications for dosing.

The team is also cultivating stem cell lines to produce organoids replicating various intestinal sections: enteroids from the jejunum and colonoids from the colon. Organoids representing other intestinal sections are also being developed. These advanced models have been shown to exhibit expression and activity of metabolic enzymes, lacking in Caco-2 cells, and transporters that will more closely mimic human intestinal functions compared to cell line-based *in vitro* models. The team has created both apical-out and basolateral-out organoid configurations for eventual use in transport and permeability assays. Preliminary observations indicate that the organoids have *in vivo*-like micromorphology, including the presence of goblet cell-like structures, limited MUC2 production, and a brush border membrane.

Initial results demonstrate that enteroids exhibit polarized expression and activity of fatty acid transport machinery, indicating a tight and polarized epithelial cell layer necessary for later transport and permeability measurements. Other transporters and metabolic enzymes present in these models are being actively characterized and compared to primary human enterocytes, with promising initial findings. Apical-out jejunal enteroids show approximately one-third the level of CYP3A4 expression found in primary human jejunal mucosa, sufficient for pharmacokinetic activity measurements. The team is also investigating enteroid protein expression using targeted proteomics to determine the correlation between activity and protein expression levels in these systems, particularly focusing on ADME-relevant proteins.

4.3.2 Development and implementation of primary human intestinal organoid models for gastrointestinal toxicity

Julia Co, a senior principal scientist at Genentech, presented her team's work toward democratizing complex *in vitro* systems for broader applications across various stages of drug discovery and development. Potential purposes include drug toxicity, ADME, biomarker discovery, screening, disease modeling, and understanding diverse patient responses, particularly in gastrointestinal toxicity contexts.

Using a 96-well plate format, the team utilized primary intestinal organoid models, developed from adult stem cells from the colon and ileum, in comparative studies. Work demonstrating their effectiveness in mimicking *in vivo* responses is underway, which could aid in the predictive evaluation of compound toxicity. This approach has increased the number of compounds that can be evaluated and has allowed the team to prioritize their library of compounds for further development and de-risk the toxicity for some lead compounds. Beyond comparative analysis of compounds, this organoid model may help to address the challenge of translating *in vitro* findings to clinical contexts, demonstrated in Belair et al. (2020) as well as by the Genentech team's work using the International Consortium for Innovation & Quality in Pharmaceutical Development's Microphysiological Systems Affiliate's (IQ-MPS)[25] recommended list of compounds for qualification of *in vitro* assays (Peters et al., 2020). Cytotoxicity

[25] https://www.iqmps.org/

in the organoid model correlated with clinical outcome, notably clinical diarrhea.

Co's team also optimized transwell monolayers using dissociated organoid cells, which allow for TEER measurements of barrier function not possible with the organoid format. The transwell system was used for evaluation of drug concentrations that cannot be reached *in vivo*, aiding evaluation during the drug development process.

Finally, Co described an innovative approach to address the labor-intensive process of scaling organoid cultures. Her team developed novel suspension methods, which they termed basement membrane extract-embedded organoid bead assemblies (BOBA), consisting of suspended droplets; and syringe-extruded organoid BME assemblies (SOBA) consisting of suspended filaments or filament fragments of cells embedded in extracellular matrix (Co et al., 2023). Both techniques allow for three-dimensional organoid growth within matrix suspended in the culture medium, overcoming the technical challenges of conventional well-based methods. Organoids produced using BOBA and SOBA maintain the essential characteristics of well-based organoids, paving the way for greater accessibility and larger-scale studies.

4.3.3 Gastrointestinal toxicity in model animal species using organ tissue equivalents

In drug development, progression into clinical trials and eventual approval is ideally best supported by data from testing in models most representative of human response. To address this, Colin Bishop, a professor at the Wake Forest Institute for Regenerative Medicine, described a project aiming at constructing functional organ tissue equivalents (OTEs) from multiple species and organs and comparing their responses to drugs with those seen using human-derived OTEs.

The procedure involves creating OTEs from commercially sourced primary cells from the liver, lung, and gut of human, mouse, rat, dog, and primate. These OTEs, which accurately recapitulate the cell proportions found in human organs, undergo screening with known drugs, biological responses are recorded, and differences are evaluated to determine the most representative model.

One significant element of the approach was the use of spheroids instead of two-dimensional transwell models, aiming for a physiologically relevant, high-throughput system. In brief, to create spheroids, the organ-specific cell types are mixed together, plated into 96- or 384-well plates, and incubated until cells compact into spheroids. Spheroids remain viable for about a month, and staining/microscopy studies indicate that organoid expression of transporters and enzymes is reflective of the *in vivo* organ.

Bishop shared findings from preliminary drug screening on organoids using six drugs: two universal toxicants, two metabolites, and two hepatotoxins. Seven days after seeding, drugs were added at a range of dilutions and viability was assessed by ATP toxicity assay three days later to obtain IC_{50} values. Results reflected clear differences between species, several of which were supported by findings from the scientific literature. Current efforts focus on improving the reproducibility of the model and screening a larger panel of compounds.

Changing gears, Bishop described collaborative work to create a biobank of gut organoids from autistic children. Given that many autistic children experience gastrointestinal issues, the ultimate goal is to investigate the gut-brain connection in autism, considering the potential role of the enteric nervous system. Despite challenges in obtaining control samples, a substantial inventory of organoids, representing a diverse spectrum of autistic profiles, has been collected to date.

5 Discussion: Strengths, limitations, and establishing scientific confidence

Following the presentations, workshop participants split into groups to discuss guiding questions related to establishing confidence in existing NAMs and the strengths and limitations of various model systems. The ensuing discussions, summarized in this section, provided a comprehensive perspective on the current state and future directions of this field, including challenges, data gaps, and critical needs.

5.1 Readiness of NAMs depends on context of use

Break-out group discussions reinforced the crucial role of GIT NAMs in three main areas: de-risking systemic toxicity and understanding absorption/pharmacokinetics; predicting GIT toxicity; and evaluating potential allergenicity. Regarding the most critical needs for future NAM development, participants emphasized the need for considering realistic exposure scenarios, such as incorporating food matrices, to better predict GIT absorption and metabolism. Discussants also noted the importance of identifying major metabolites and considering the potential impact of gastric residence time in systemic toxicity. Additionally, discussions indicated the need to recapitulate the role of the microbiome, possibly through synthetic biology. Participants also mentioned that comprehensive morphological evaluations of models are a useful aspect of characterization. Lastly, participants noted that providing a clear explanation of regulatory needs and contexts of use could help to avoid an overwhelmingly broad research focus.

Participants noted that the readiness of GIT-related NAMs for regulatory use and other forms of GIT-related testing varies across models and applications. The workshop presentations and the discussion sessions highlighted the availability of NAMs that can assess overt toxicity and barrier disruption from acute exposures. Routinely used Caco-2 models address direct cytotoxicity; and TEER can be effectively employed as a relatively high-throughput assessment method to assess barrier integrity. Permeability assays have also proven useful but exhibit lower throughput. Systems that incorporate flow dynamics are also showing promise in enhancing the biological accuracy of *in vitro* models.

Discussants noted that three-dimensional NAMs, such as intestinal organoids, may be particularly well-suited for evaluating mechanistic aspects of GIT function and toxicity, based on the tunability of organoids to various cell populations via culture conditions. Organoids and other NAMs derived from primary human cells were acknowledged for their capability to recapitulate human diversity. Establishment of diverse biobanks of intestinal orga-

noids could facilitate studying specific susceptible populations. Organoid systems are also showing promise in the evaluation of potential species-specific toxicity.

The maturity of NAMs relies on various factors, including the establishment of standardized approaches that generate reproducible results. Reproducibility should also include understanding and controlling for sources of variability, and differentiating between technical versus biological variability (i.e., differences among people within a population). Other criteria supporting model maturation include demonstrated biological relevance, quality assurance, and transparency (e.g., OECD's Guidance Documents on Good In Vitro Method Practices, including Good Cell Culture Practice (OECD, 2018)), and the establishment of curated sets of reference compounds for method evaluation. Factors such as the type of substance being assessed, required throughput, and practical constraints of the assay (e.g., continued commercial availability, ease of use, transferability) must also be considered.

In response to a guiding question about successful examples of using NAMs for any purpose, participants shared several successful outcomes of using NAMs, such as the successful testing of a synthetic microbe for the treatment of phenylketonuria (see 4.2.3), and the use of patient-derived organoids to expand indications of a cystic fibrosis drug (Conti et al., 2022; De Poel et al., 2023). Several *in silico* NAMs have demonstrated success in ADME studies, including Certara's Simcyp™ PBPK simulator[26] and the Collaborative Acute Toxicity Modeling Suite (CATMoS) (Mansouri et al., 2021). Despite these successes, ongoing challenges include recapitulating long-term toxicity, understanding the role of the inflammatory/immune response, and developing well-defined models to characterize metabolic pathway perturbations. In addition, assessing volatiles and nanomaterials were noted as particularly challenging.

In summary, discussions emphasized that defining the COU for each model is crucial for establishing criteria for model maturity and readiness for regulatory or research applications. Multiple participants stressed the need to avoid the "one-model solution" approach, recognizing that no single model will be perfect for all endpoints; instead, multiple models will likely be necessary. The choice between simpler two-dimensional models and more complex three-dimensional or perfusion-based systems depends on the specific application and the level of physiological accuracy required by the scientific question and consequent regulatory needs. For example, while two-dimensional models may be adequate for investigating acute toxicity, three-dimensional systems like organoids may be needed for long-term studies (e.g., tissue remodeling) and for capturing human diversity. Striking a balance between complexity and practicality/ease of use is crucial for NAMs to be effective and widely adoptable.

5.2 Reference data for NAMs

Participants were asked to identify the "gold standard" against which NAMs should be compared to establish confidence. For GIT toxicity, clinical human data can serve as appropriate reference data when available (e.g., bioavailability data for absorption). Pharmacokinetic modeling could also be applied to understand dose relationship to internal concentration and exposure. For example, the IQ-MPS[25] has assembled assay qualification compound sets for nausea, vomiting, constipation, and diarrhea (Peters et. al, 2020). These lists were derived based on clinical data for approved drugs, and they may provide valuable data for assessing the predictivity of *in vitro* models. However, not all chemical classes, mechanisms of action, and potencies may be represented as these lists only include approved human pharmaceuticals.

Although several participants voiced the opinion that NAMs may be better suited for assessing various aspects of GIT toxicity compared to many animal-based *in vivo* methods, it was acknowledged that, in some cases, animals have demonstrated high predictivity of GIT toxicity (e.g., non-human primates). When such *in vivo* animal data are available, they can be valuable in assessing the performance of an *in vitro* method. Developers of *in vitro* GIT models have noted that animal data remain important in understanding IVIVE, and the development of *in vitro* models with animal cells can be warranted in specific scenarios (Belair et al., 2020; Peters et al., 2020).

Validating to animal models can be problematic in certain cases. For example, animal species used in toxicity studies have different susceptibilities to diarrhea. Predictive *in vitro* assays to assess drug induction of diarrhea in humans could be useful in screening potential drug candidates prior to further testing, such as those using human ileal organoids (Belair et al., 2020). However, additional work is needed to determine how *in vitro* drug concentrations in this model translate to *in vivo* exposures. Development of similar *in vitro* models with animal-derived cells has also been suggested as a way to validate the *in vivo* predictivity of the model and to enable a better understanding of species-specific findings. Participants also discussed whether bridging assessments might be appropriate to compare animal-based microphysiological systems to animal data, providing confidence in human cell-based systems.

5.3 Remaining challenges and data gaps

Discussions highlighted significant data gaps that hinder the routine application of NAMs. A primary concern is the lack of appropriate statistical characterization and data interpretation. Much of the NAM data were noted to be of low confidence, often due to issues with calibration chemicals or control measurements (e.g., positive controls (Petersen et al., 2021)). Other statistical considerations, such as determining appropriate sample sizes, are also necessary to provide NAMs with credibility in toxicological assessments. Participants also mentioned metabolism-related data gaps, noting that GIT differences between animals and humans are difficult to model *in vitro* and that multiple models are needed to capture the diversity of the human population or the varying density of the mucous layer. Further, simple models that do not recapitulate metabolism may not be "tissue-like" enough to use for decision making. Finally, the discussions underscored the ne-

[26] https://www.certara.com/software/simcyp-pbpk/

cessity for standardization that aligns with regulatory demands and involves detailed characterization of model tissues while also considering cost-effectiveness.

Participants discussed key functionalities that require prioritization for advancing NAM development for regulatory use. Barrier function, metabolism, and absorption were specifically noted as important priorities. While barrier integrity was considered mostly "solved" (e.g., TEER assay), the scaleup of primary cells for use in barrier function assays was considered a lingering challenge. Addressing metabolism-related gaps involves creating organoids from primary cells and establishing models of first-pass metabolism. Intestinal efflux transporters were also mentioned as an important functionality, but participants predicted that this function might naturally emerge if barrier function and metabolism are addressed. Finally, while multi-organ models are beginning to emerge to better recapitulate human-relevant functionality, discussants emphasized the need to balance the complexity of those models against the practicality and specificity of simpler models.

To enhance the relevance and applicability of GIT NAMs in pharmaceutical and regulatory contexts, these models must closely mimic the human GIT, both structurally and functionally. Participants noted that various regions of the GIT, such as the small and large intestines, require distinct modeling approaches to accurately recapitulate region-specific tissue architecture (e.g., intestinal villi, tight junctions, brush border membrane, mucous layer). Further, participants indicated a need for whole-gut representation, beginning with the oral cavity and assuring proper gastric residence time, to reflect the diverse chemical transformations occurring throughout the digestive process. For instance, multiple-organ models that are conditioned with GIT fluids and mimic the role of the liver in first-pass metabolism might more closely approximate human responses. Furthermore, integrating microbiome diversity was deemed pivotal by participants, particularly alterations to the microbiome characteristic of some disease states. Modeling the anaerobic environment of the GIT was acknowledged as a significant challenge. Finally, despite progress, NAMs still cannot accurately assess long-term toxicity, inflammatory or immune responses, nor nausea and vomiting (Holmes et al., 2009), which frequently affect medication compliance. Addressing these challenges will not only enhance the reliability of NAMs in predicting human responses but also aid in their acceptance in regulatory contexts.

References

Artursson, P., Palm, K. and Luthman, K. (2001). Caco-2 monolayers in experimental and theoretical predictions of drug transport. 1PII of original article: S0169-409X(96)00415-2. The article was originally published in *Adv Drug Deliv Rev 22* (1996), 67-84.1. *Adv Drug Del Rev 46*, 27-43. doi:10.1016/S0169-409X(00)00128-9

Atkins, J. T., George, G. C., Hess, K. et al. (2020). Pre-clinical animal models are poor predictors of human toxicities in phase 1 oncology clinical trials. *Br J Cancer 123*, 1496-1501. doi:10.1038/s41416-020-01033-x

Avila, A. M., Bebenek, I., Bonzo, J. A. et al. (2020). An FDA/CDER perspective on nonclinical testing strategies: Classical toxicology approaches and new approach methodologies (NAMs). *Regul Toxicol Pharmacol 114*, 104662. doi:10.1016/j.yrtph.2020.104662

Avila, A. M., Bebenek, I., Mendrick, D. L. et al. (2023). Gaps and challenges in nonclinical assessments of pharmaceuticals: An FDA/CDER perspective on considerations for development of new approach methodologies. *Regul Toxicol Pharmacol 139*, 105345. doi:10.1016/j.yrtph.2023.105345

Bein, A., Shin, W., Jalili-Firoozinezhad, S. et al. (2018). Microfluidic organ-on-a-chip models of human intestine. *Cell Mol Gastroenterol Hepatol 5*, 659-668. doi:10.1016/j.jcmgh.2017.12.010

Belair, D. G., Visconti, R. J., Hong, M. et al. (2020). Human ileal organoid model recapitulates clinical incidence of diarrhea associated with small molecule drugs. *Toxicol In Vitro 68*, 104928. doi:10.1016/j.tiv.2020.104928

Bell, S. M., Chang, X., Wambaugh, J. F. et al. (2018). In vitro to in vivo extrapolation for high throughput prioritization and decision making. *Toxicol In Vitro 47*, 213-227. doi:10.1016/j.tiv.2017.11.016

Blutt, S. E., Broughman, J. R., Zou, W. et al. (2017). Gastrointestinal microphysiological systems. *Exp Biol Med (Maywood) 242*, 1633-1642. doi:10.1177/1535370217710638

Bøgh, K. L., Van Bilsen, J., Głogowski, R. et al. (2016). Current challenges facing the assessment of the allergenic capacity of food allergens in animal models. *Clin Transl Allergy 6*, 21. doi:10.1186/s13601-016-0110-2

Bova, R. A., Lamont, A. C., Picou, T. J. et al. (2023). Shiga toxin (Stx) type 1a and Stx2a translocate through a three-layer intestinal model. *Toxins 15*, 207. doi:10.3390/toxins15030207

Calero-Medina, L., Jimenez-Casquet, M. J., Heras-Gonzalez, L. et al. (2023). Dietary exposure to endocrine disruptors in gut microbiota: A systematic review. *Sci Total Environ 886*, 163991. doi:10.1016/j.scitotenv.2023.163991

Cattaneo, I., Astuto, M. C., Binaglia, M. et al. (2023). Implementing new approach methodologies (NAMs) in food safety assessments: Strategic objectives and actions taken by the European Food Safety Authority. *Trends Food Sci Technol 133*, 277-290. doi:10.1016/j.tifs.2023.02.006

Chang, X., Tan, Y.-M., Allen, D. G. et al. (2022). IVIVE: Facilitating the use of in vitro toxicity data in risk assessment and decision making. *Toxics 10*, 232. doi:10.3390/toxics10050232

Chi, D. S., DeKruyff, R. H., Lerman, S. P. et al. (1982). Effect of histoincompatibility on migration of labeled bursa cells in the chicken. *Transplantation 33*, 377-381. doi:10.1097/00007890-198204000-00007

Chiu, K., Warner, G., Nowak, R. A. et al. (2020). The impact of environmental chemicals on the gut microbiome. *Toxicol Sci 176*, 253-284. doi:10.1093/toxsci/kfaa065

Co, J. Y., Klein, J. A., Kang, S. et al. (2023). Suspended hydrogel culture as a method to scale up intestinal organoids. *Sci Rep 13*, 10412. doi:10.1038/s41598-023-35657-9

Coecke, S., Pelkonen, O., Leite, S. B. et al. (2013). Toxicokinetics as a key to the integrated toxicity risk assessment based primar-

ily on non-animal approaches. *Toxicol In Vitro 27*, 1570-1577. doi:10.1016/j.tiv.2012.06.012

Collins, F. S., Gray, G. M. and Bucher, J. R. (2008). Transforming environmental health protection. *Science 319*, 906-907. doi:10.1126/science.1154619

Conti, J., Sorio, C. and Melotti, P. (2022). Organoid technology and its role for theratyping applications in cystic fibrosis. *Children (Basel) 10*, 4. doi:10.3390/children10010004

Cook, D., Brown, D., Alexander, R. et al. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nat Rev Drug Discov 13*, 419-431. doi:10.1038/nrd4309

Costa, J. and Ahluwalia, A. (2019). Advances and current challenges in intestinal in vitro model engineering: A digest. *Front Bioeng Biotechnol 7*, 144. doi:10.3389/fbioe.2019.00144

Cote, I., Andersen, M. E., Ankley, G. T. et al. (2016). The next generation of risk assessment multi-year study – Highlights of findings, applications to risk assessment, and future directions. *Environ Health Perspect 124*, 1671-1682. doi:10.1289/ehp233

Crevel, R. W. R., Baumert, J. L., Baka, A. et al. (2014). Development and evolution of risk assessment for food allergens. *Food Chem Toxicol 67*, 262-276. doi:10.1016/j.fct.2014.01.032

Croote, D., Darmanis, S., Nadeau, K. C. et al. (2018). High-affinity allergen-specific human antibodies cloned from single IgE B cell transcriptomes. *Science 362*, 1306-1309. doi:10.1126/science.aau2599

Croote, D., Wong, J. J. W., Pecalvel, C. et al. (2024). Widespread monoclonal IgE antibody convergence to an immunodominant, proanaphylactic Ara h 2 epitope in peanut allergy. *J Allergy Clin Immunol 153*, 182-192.e7. doi:10.1016/j.jaci.2023.08.035

Dahlgren, D., Roos, C., Sjögren, E. et al. (2015). Direct in vivo human intestinal permeability (Peff) determined with different clinical perfusion and intubation methods. *J Pharm Sci 104*, 2702-2726. doi:10.1002/jps.24258

De Castelbajac, T., Aiello, K., Arenas, C. G. et al. (2023). Innovative tools and methods for toxicity testing within PARC work package 5 on hazard assessment. *Front Toxicol 5*, 1216369. doi:10.3389/ftox.2023.1216369

Dearman, R. J., Basketter, D. A. and Kimber, I. (2013). Inter-relationships between different classes of chemical allergens. *J Appl Toxicol 33*, 558-565. doi:10.1002/jat.1758

De Poel, E., Spelier, S., Hagemeijer, M. C. et al. (2023). FDA-approved drug screening in patient-derived organoids demonstrates potential of drug repurposing for rare cystic fibrosis genotypes. *J Cyst Fibros 22*, 548-559. doi:10.1016/j.jcf.2023.03.004

DeLoid, G. M., Wang, Y., Kapronezai, K. et al. (2017). An integrated methodology for assessing the impact of food matrix and gastrointestinal effects on the biokinetics and cellular toxicity of ingested engineered nanomaterials. *Part Fibre Toxicol 14*, 40. doi:10.1186/s12989-017-0221-5

Drenckhahn, D. and Dermietzel, R. (1988). Organization of the actin filament cytoskeleton in the intestinal brush border: A quantitative and qualitative immunoelectron microscope study. *J Cell Biol 107*, 1037-1048. doi:10.1083/jcb.107.3.1037

Ede, J. D., Ong, K. J., Mulenos, M. R. et al. (2020). Physical, chemical, and toxicological characterization of sulfated cellulose nanocrystals for food-related applications using in vivo and in vitro strategies. *Toxicol Res (Camb) 9*, 808-822. doi:10.1093/toxres/tfaa082

EFSA NDA Panel (EFSA Panel on Dietetic Products, Nutrition and Allergies), Turck, D., Bresson, J.-L. et al. (2021). Guidance on the preparation and submission of an application for authorisation of a novel food in the context of Regulation (EU) 2015/2283 (Revision 1). *EFSA J 19*, e06555. doi:10.2903/j.efsa.2021.6555

EFSA Scientific Committee, More, S., Bampidis, V. et al. (2021a). Guidance on risk assessment of nanomaterials to be applied in the food and feed chain: Human and animal health. *EFSA J 19*, e06768. doi:10.2903/j.efsa.2021.6768

EFSA Scientific Committee, More, S. J., Bampidis, V. et al. (2021b). Guidance document on scientific criteria for grouping chemicals into assessment groups for human risk assessment of combined exposure to multiple chemicals. *EFSA J 19*, e07033. doi:10.2903/j.efsa.2021.7033

Falcón-Cano, G., Molina, C. and Cabrera-Pérez, M. Á. (2022). Reliable prediction of Caco-2 permeability by supervised recursive machine learning approaches. *Pharmaceutics 14*, 1998. doi:10.3390/pharmaceutics14101998

Federer, C., Yoo, M. and Tan, A. C. (2016). Big data mining and adverse event pattern analysis in clinical drug trials. *Assay Drug Dev Technol 14*, 557-566. doi:10.1089/adt.2016.742

Fedi, A., Vitale, C., Ponschin, G. et al. (2021). In vitro models replicating the human intestinal epithelium for absorption and metabolism studies: A systematic review. *J Control Release 335*, 247-268. doi:10.1016/j.jconrel.2021.05.028

Fenwick, N., Griffin, G. and Gauthier, C. (2009). The welfare of animals used in science: How the "Three Rs" ethic guides improvements. *Can Vet J 50*, 523-530.

Fogh, J., Fogh, J. M. and Orfeo, T. (1977). One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. *J Natl Cancer Inst 59*, 221-226. doi:10.1093/jnci/59.1.221

Gelberg, H. (2018). Pathophysiological mechanisms of gastrointestinal toxicity. In *Comprehensive Toxicology* (139-178). Elsevier. doi:10.1016/B978-0-12-801238-3.10923-7

Gibb, M., Pradhan, S. H., Mulenos, M. R. et al. (2021). Characterization of a human in vitro intestinal model for the hazard assessment of nanomaterials used in cancer immunotherapy. *Appl Sci 11*, 2113. doi:10.3390/app11052113

Gustafsson, J. K., Davis, J. E., Rappai, T. et al. (2021). Intestinal goblet cells sample and deliver lumenal antigens by regulated endocytic uptake and transcytosis. *Elife 10*, e67292. doi:10.7554/elife.67292

Hoffmann, S., Aiassa, E., Angrish, M. et al. (2022). Application of evidence-based methods to construct mechanism-driven chemical assessment frameworks. *ALTEX 39*, 499-518. doi:10.14573/altex.2202141

Hoh, R. A., Joshi, S. A., Lee, J.-Y. et al. (2020). Origins and clonal convergence of gastrointestinal IgE^+ B cells in human peanut allergy. *Sci Immunol 5*, eaay4209. doi:10.1126/sciimmunol.aay4209

Holmes, A., Rudd, J., Tattersall, F. et al. (2009). Opportunities for the replacement of animals in the study of nausea and vomiting. *Br J Pharmacol 157*, 865-880. doi:10.1111/j.1476-5381.2009.00176.x

Honda, G, Kenyon, E. M., Davidson-Fritz, S. et al. (2024). Impact of gut permeability on estimation of oral bioavailability for chemicals in commerce and the environment. *Toxicol In Vitro*, submitted.

Jamei, M., Marciniak, S., Feng, K. et al. (2009). The Simcyp® population-based ADME simulator. *Expert Opin Drug Metab Toxicol 5*, 211-223. doi:10.1517/17425250802691074

Kang, T. and Kim, H. (2016). Farewell to animal testing: Innovations on human intestinal microphysiological systems. *Micromachines 7*, 107. doi:10.3390/mi7070107

Kim, M. T., Sedykh, A., Chakravarti, S. K. et al. (2014). Critical evaluation of human oral bioavailability for pharmaceutical drugs by using various cheminformatics approaches. *Pharm Res 31*, 1002-1014. doi:10.1007/s11095-013-1222-1

Kavlock, R. J., Bahadori, T., Barton-Maclaren, T. S. et al. (2018). Accelerating the pace of chemical risk assessment. *Chem Res Toxicol 31*, 287-290. doi:10.1021/acs.chemrestox.7b00339

Kazemi, S., Danisman, E. and Epstein, M. M. (2023). Animal models for the study of food allergies. *Curr Protoc 3*, e685. doi:10.1002/cpz1.685

Kleinstreuer, N. C., Karmaus, A. L., Mansouri, K. et al. (2018). Predictive models for acute oral systemic toxicity: A workshop to bridge the gap from research to regulation. *Comput Toxicol 8*, 21-24. doi:10.1016/j.comtox.2018.08.002

Kopec, A. K., Yokokawa, R., Khan, N. et al. (2021). Microphysiological systems in early stage drug development: Perspectives on current applications and future impact. *J Toxicol Sci 46*, 99-114. doi:10.2131/jts.46.99

Krewski, D., Andersen, M. E., Tyshenko, M. G. et al. (2020). Toxicity testing in the 21st century: Progress in the past decade and future perspectives. *Arch Toxicol 94*, 1-58. doi:10.1007/s00204-019-02613-4

Kulkarni, D. H., Gustafsson, J. K., Knoop, K. A. et al. (2020). Goblet cell associated antigen passages support the induction and maintenance of oral tolerance. *Mucosal Immunol 13*, 271-282. doi:10.1038/s41385-019-0240-7

Lee, B., Moon, K. M. and Kim, C. Y. (2018). Tight junction in the intestinal epithelium: Its association with diseases and regulation by phytochemicals. *J Immunol Res 2018*, 2645465. doi:10.1155/2018/2645465

Licari, A., Votto, M., D'Auria, E. et al. (2020). Eosinophilic gastrointestinal diseases in children: A practical review. *Curr Pediatr Rev 16*, 106-114. doi:10.2174/1573396315666191022154432

Lopez-Escalera, S. and Wellejus, A. (2022). Evaluation of Caco-2 and human intestinal epithelial cells as in vitro models of colonic and small intestinal integrity. *Biochem Biophys Rep 31*, 101314. doi:10.1016/j.bbrep.2022.101314

Manatakis, D. V., VanDevender, A. and Manolakos, E. S. (2021). An information-theoretic approach for measuring the distance of organ tissue samples using their transcriptomic signatures. *Bioinformatics 36*, 5194-5204. doi:10.1093/bioinformatics/btaa654

Mansouri, K., Karmaus, A. L., Fitzpatrick, J. et al. (2021). CATMoS: Collaborative acute toxicity modeling suite. *Environ Health Perspect 129*, 47013. doi:10.1289/ehp8495

Marrero, D., Pujol-Vila, F., Vera, D. et al. (2021). Gut-on-a-chip: Mimicking and monitoring the human intestine. *Biosens Bioelectron 181*, 113156. doi:10.1016/j.bios.2021.113156

Marshall, L. J., Bailey, J., Cassotta, M. et al. (2023). Poor translatability of biomedical research using animals – A narrative review. *Altern Lab Anim 51*, 102-135. doi:10.1177/02611929231157756

Martin, O. V. (2023). Synergistic effects of chemical mixtures: How frequent is rare? *Curr Opin Toxicol 36*, 100424. doi:10.1016/j.cotox.2023.100424

Martin-Folgar, R., González-Caballero, M. C., Torres-Ruiz, M. et al. (2024). Molecular effects of polystyrene nanoplastics on human neural stem cells. *PLoS One 19*, e0295816. doi:10.1371/journal.pone.0295816

Marx-Stoelting, P., Rivière, G., Luijten, M. et al. (2023). A walk in the PARC: Developing and implementing 21st century chemical risk assessment in Europe. *Arch Toxicol 97*, 893-908. doi:10.1007/s00204-022-03435-7

McDole, J. R., Wheeler, L. W., McDonald, K. G. et al. (2012). Goblet cells deliver luminal antigen to CD103+ dendritic cells in the small intestine. *Nature 483*, 345-349. doi:10.1038/nature10863

Mestas, J. and Hughes, C. C. W. (2004). Of mice and not men: Differences between mouse and human immunology. *J Immunol 172*, 2731-2738. doi:10.4049/jimmunol.172.5.2731

Min, J., Keswani, T., LaHood, N. A. et al. (2024). Design of an Ara h 2 hypoallergen from conformational epitopes. *Clin Exp Allergy 54*, 46-55. doi:10.1111/cea.14433

Minekus, M., Alminger, M., Alvito, P. et al. (2014). A standardised static in vitro digestion method suitable for food – An international consensus. *Food Funct 5*, 1113-1124. doi:10.1039/c3fo60702j

Monticello, T. M., Jones, T. W., Dambach, D. M. et al. (2017). Current nonclinical testing paradigm enables safe entry to first-in-human clinical trials: The IQ consortium nonclinical to clinical translational database. *Toxicol Appl Pharmacol 334*, 100-109. doi:10.1016/j.taap.2017.09.006

Musther, H., Olivares-Morales, A., Hatley, O. J. D. et al. (2014). Animal versus human oral drug bioavailability: Do they correlate? *Eur J Pharm Sci 57*, 280-291. doi:10.1016/j.ejps.2013.08.018

NRC – National Research Council (1983). Risk Assessment in the Federal Government: Managing the Process. Washington, D.C., USA: National Academies Press. doi:10.17226/366

Nelson, M. T., Charbonneau, M. R., Coia, H. G. et al. (2021). Characterization of an engineered live bacterial therapeutic for the treatment of phenylketonuria in a human gut-on-a-chip. *Nat Commun 12*, 2805. doi:10.1038/s41467-021-23072-5

Nelson, M. T., Coia, H. G., Holt, C. et al. (2023). Evaluation of human performance aiding live synthetically engineered bacteria in a gut-on-a-chip. *ACS Biomater Sci Eng 9*, 5136-5150. doi:10.1021/acsbiomaterials.2c00774

Newberry, R. D. and Hogan, S. P. (2021). Intestinal epithelial cells in tolerance and allergy to dietary antigens. *J Allergy Clin Immunol 147*, 45-48. doi:10.1016/j.jaci.2020.10.030

Noah, T. K., Lee, J.-B., Brown, C. A. et al. (2021). Thermoneutrality alters gastrointestinal antigen passage patterning and predisposes to oral antigen sensitization in mice. *Front Immunol 12*, 636198. doi:10.3389/fimmu.2021.636198

OECD (2018). Guidance Document on Good In Vitro Method Practices (GIVIMP). *OECD Series on Testing and Assessment, No. 286*. OECD Publishing, Paris. doi:10.1787/9789264304796-en

Ölander, M., Wiśniewski, J. R., Matsson, P. et al. (2016). The proteome of filter-grown Caco-2 cells with a focus on proteins involved in drug disposition. *J Pharm Sci 105*, 817-827. doi:10.1016/j.xphs.2015.10.030

O'Reilly, M., Mellotte, G., Ryan, B. et al. (2020). Gastrointestinal side effects of cancer treatments. *Ther Adv Chronic Dis 11*, 204062232097035. doi:10.1177/2040622320970354

Pali-Schöll, I., Verhoeckx, K., Mafra, I. et al. (2019). Allergenic and novel food proteins: State of the art and challenges in the allergenicity assessment. *Trends Food Sci Technol 84*, 45-48. doi:10.1016/j.tifs.2018.03.007

Pallocca, G., Moné, M. J., Kamp, H. et al. (2022). Next-generation risk assessment of chemicals – Rolling out a human-centric testing strategy to drive 3R implementation: The RISK-HUNT3R project perspective. *ALTEX 39*, 419-426. doi:10.14573/altex.2204051

Patil, S. U., Ogunniyi, A. O., Calatroni, A. et al. (2015). Peanut oral immunotherapy transiently expands circulating Ara h 2-specific B cells with a homologous repertoire in unrelated subjects. *J Allergy Clin Immunol 136*, 125-134.e12. doi:10.1016/j.jaci.2015.03.026

Peters, M. F., Choy, A. L., Pin, C. et al. (2020). Developing in vitro assays to transform gastrointestinal safety assessment: Potential for microphysiological systems. *Lab Chip 20*, 1177-1190. doi:10.1039/c9lc01107b

Petersen, E. J., Nguyen, A., Brown, J. et al. (2021). Characteristics to consider when selecting a positive control material for an in vitro assay. *ALTEX 38*, 365-376. doi:10.14573/altex.2102111

Petersen, E. J., Ceger, P., Allen, D. G. et al. (2022). U.S. Federal Agency interests and key considerations for new approach methodologies for nanomaterials. *ALTEX 39*, 183-206. doi:10.14573/altex.2105041

Petersen, E. J., Elliott, J. T., Gordon, J. et al. (2023). Technical framework for enabling high-quality measurements in new approach methodologies (NAMs). *ALTEX 40, 174-186*. doi:10.14573/altex.2205081

Pirmohamed, M. (2023). Pharmacogenomics: Current status and future perspectives. *Nat Rev Genet 24*, 350-362. doi:10.1038/s41576-022-00572-8

Pradhan, S. H., Mulenos, M. R., Steele, L. R. et al. (2020). Physical, chemical, and toxicological characterization of fibrillated forms of cellulose using an in vitro gastrointestinal digestion and co-culture model. *Toxicol Res (Camb) 9*, 290-301. doi:10.1093/toxres/tfaa026.

Price, G. and Patel, D. A. (2024). Drug bioavailability. In *StatPearls*. Treasure Island, FL, USA: StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/nbk557852/ (accessed 02.06.2024)

Punt, A., Louisse, J., Pinckaers, N. et al. (2022). Predictive performance of next generation physiologically based kinetic (PBK) model predictions in rats based on in vitro and in silico input data. *Toxicol Sci 186*, 18-28. doi:10.1093/toxsci/kfab150

Qureshi, Z. P., Seoane-Vazquez, E., Rodriguez-Monguio, R. et al. (2011). Market withdrawal of new molecular entities approved in the United States from 1980 to 2009. *Pharmacoepidemiol Drug Saf 20*, 772-777. doi:10.1002/pds.2155

Richard, A. M., Judson, R. S., Houck, K. A. et al. (2016). ToxCast chemical landscape: Paving the road to 21st century toxicology. *Chem Res Toxicol 29*, 1225-1251. doi:10.1021/acs.chemrestox.6b00135

Rotroff, D. M., Wetmore, B. A., Dix, D. J. et al. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci 117*, 348-358. doi:10.1093/toxsci/kfq220

Sayes, C. M., Reed, K. L. and Warheit, D. B. (2007). Assessing toxicity of fine and nanoparticles: Comparing in vitro measurements to in vivo pulmonary toxicity profiles. *Toxicol Scie 97*, 163-180. doi:10.1093/toxsci/kfm018

Tarazona, J. V., Cattaneo, I., Niemann, L. et al. (2022). A tiered approach for assessing individual and combined risk of pyrethroids using human biomonitoring data. *Toxics 10*, 451. doi:10.3390/toxics10080451

Thomas, R. S., Philbert, M. A., Auerbach, S. S. et al. (2013). Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicol Sci 136*, 4-18. doi:10.1093/toxsci/kft178

Thompson, C. L., Fu, S., Knight, M. M. et al. (2020). Mechanical stimulation: A crucial element of organ-on-chip models. *Front Bioeng Biotechnol 8*, 602646. doi:10.3389/fbioe.2020.602646

Tsaioun, K. (2016). Evidence-based absorption, distribution, metabolism, excretion (ADME) and its interplay with alternative toxicity methods. *ALTEX 33*, 343-358. doi:10.14573/altex.1610101

Turner, J. (2023). Incorporating new approach methodologies into regulatory nonclinical pharmaceutical safety assessment. *ALTEX 40*, 519-533. doi:10.14573/altex.2212081

Valiei, A., Aminian-Dehkordi, J. and Mofrad, M. R. K. (2023). Gut-on-a-chip models for dissecting the gut microbiology and physiology. *APL Bioeng 7*, 011502. doi:10.1063/5.0126541

Van Norman, G. A. (2019). Limitations of animal studies for predicting toxicity in clinical trials: Is it time to rethink our current approach? *JACC Basic Transl Sci 4*, 845-854. doi:10.1016/j.jacbts.2019.10.008

Ventola, C. L. (2013). Role of pharmacogenomic biomarkers in predicting and improving drug response: Part 1: The clinical significance of pharmacogenetic variants. *P T 38*, 545-560.

Wambaugh, J. F., Hughes, M. F., Ring, C. L. et al. (2018). Evaluating in vitro-in vivo extrapolation of toxicokinetics. *Toxicol Sci 163*, 152-169. doi:10.1093/toxsci/kfy020

Wambaugh, J. F., Wetmore, B. A., Ring, C. L. et al. (2019). Assessing toxicokinetic uncertainty and variability in risk prioritization. *Toxicol Sci 172*, 235-251. doi:10.1093/toxsci/kfz205

Wang, H., Brown, P. C., Chow, E. C. Y. et al. (2021). 3D cell culture models: Drug pharmacokinetics, safety assessment, and regulatory consideration. *Clin Transl Sci 14*, 1659-1680. doi:10.1111/cts.13066

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S. et al. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci 125*, 157-174. doi:10.1093/toxsci/kfr254

Wetmore, B. A., Wambaugh, J. F., Allen, B. et al. (2015). Incorporating high-throughput exposure predictions with dosimetry-adjusted in vitro bioactivity to inform chemical toxicity testing. *Toxicol Sci 148*, 121-136. doi:10.1093/toxsci/kfv171

Xiang, Y., Wen, H., Yu, Y. et al. (2020). Gut-on-chip: Recreating human intestine in vitro. *J Tissue Eng 11*, 204173142096531. doi:10.1177/2041731420965318

Yim, D.-S., Choi, S. and Bae, S. H. (2020). Predicting human pharmacokinetics from preclinical data: Absorption. *Transl Clin Pharmacol 28*, 126. doi:10.12793/tcp.2020.28.e14

Disclaimer

The views and work products expressed in this article are those of the authors and do not reflect the official views, policies, or positions of their institutions of employment. The statements, opinions, or conclusions contained therein were not approved by, and do not necessarily represent, statements, opinions, or conclusions of the United States government, or of any of the organizations to which the authors of this publication are affiliated. Inotiv staff provide technical support for NICEATM, but do not represent NIEHS, NTP, or any U.S. agency. Certain commercial products or equipment are described in this paper to adequately specify the experimental procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, National Institute of Environmental Health Sciences (NIEHS), the European Food Safety Authority, other governmental agencies, or other organizations, nor does it imply that it is necessarily the best available for the purpose. The text reflects the presentations and discussions and is not a consensus report.

Conflict of interest

Some authors are employed by pharmaceutical companies, as per their affiliations. SPH receives grant support in part from Regeneron Pharmaceuticals and is a consultant with Abbvie.

Acknowledgements

The project was funded in part with federal funds from the National Institute of Environmental Health Sciences, National Institute of Health, under contract no. HHSN273201500010C to Integrated Laboratory Systems LLC, an Inotiv company, and by National Institutes of Health Intramural Research Project ES103386-01, Research Operations Supporting the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods. PL has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 964537 (RISK-HUNT3R), which is part of the ASPIS cluster. Section 4.2.1 was supported in part by appointing YJ to the Research Participation Program at the U.S. FDA administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. FDA.

Acknowledgement

We would like to thank Jaehyun Lim for his participation in the workshop.

Data availability

No datasets were created for this workshop.