

## Opinion Article

# The Value of Organs-on-Chip for Regulatory Safety Assessment

Minne B. Heringa<sup>1,2</sup>, Margriet V. D. Z. Park<sup>1</sup>, Anne S. Kienhuis<sup>1</sup> and Rob J. Vandebriel<sup>1</sup>

<sup>1</sup>National Institute of Public Health and the Environment, Bilthoven, The Netherlands; <sup>2</sup>present address: RB Home and Hygiene, Amsterdam, The Netherlands

### Abstract

Organs-on-chip (OC) have gained much interest as animal-free toxicity testing methods due to their closer resemblance to human tissues and longer culture viability than conventional *in vitro* methods. The current paper discusses where and how OCs may take a role in the transition to a more predictive, animal-free safety assessment for regulatory purposes. From a preliminary analysis of a repeated dose toxicity database, ten organs of priority for OC development for regulatory use have been identified. For a number of these organs (lung, skin, liver, kidney, heart, and intestine), OCs are already at rather advanced stages of development, such that involvement of regulators becomes of value in the optimization towards fitness-for-purpose of these methods. For organs such as testis, spleen, brain, and stomach, OCs are much more premature, if existing at all. Therefore, developmental work on OCs for these latter organs is expected to stay in the academic arena for the coming time. A number of technical recommendations and some challenges to reaching final implementation are discussed. We recommend that the development of OCs goes forward together with the development of adverse outcome pathways (AOP) and that they are combined with other methods into integrated testing strategies. Overall, opportunities exist, but much still needs to be done. In our view, regular interactions in multi-stakeholder workshops on the application of animal-free innovations such as OCs will be beneficial.

## 1 Introduction

Traditionally, much of the data needed for the regulatory safety assessment of chemical substances and pharmaceuticals comes from experimental animal studies. In fact, in Europe approximately 9% of the total number of test animals is used for safety assessment (EC, 2013). However, over the last decades it has become increasingly apparent that animal studies may have limited predictive value for various toxic effects in humans (e.g., Adler et al., 2011; Arrowsmith and Miller, 2013). To overcome this, regulatory safety assessments should make much more use of the increased knowledge on the mechanisms underlying toxic effects that has been gained since the era during which today's regulatory toxicity studies were devised. Besides this scientific argument, ethical and economic arguments further challenge the use of laboratory animals.

Government, industry and academia's collaborative investment in replacement, refinement and reduction of animal experiments (3Rs) has already built a toolbox with various animal-free toxicity testing methods for use in regulatory safety assessment, but a

complete replacement of animal tests is not yet possible. Especially in case of the more complex systemic endpoints (e.g., carcinogenicity and reproductive toxicity), adequate animal-free methods are not yet available. Similarly, sufficiently predictive animal-free models to obtain information on toxicokinetics (absorption, metabolism and excretion), which is indispensable to calculate safe exposure limits for humans from *in vitro* assays, are still lacking. In short, from a regulatory point of view, the transition to a more predictive safety assessment of chemical and pharmaceutical substances requires the development of novel models and methods that reflect complex human physiology and biology *in vitro*.

In this respect, one of the fastest-growing, most promising innovative technologies is the organ-on-chip (OC) technology (Marx et al., 2016). However, its value for regulatory safety assessment of chemical substances and pharmaceuticals is still unclear. Its role is in part dependent on the future development of regulatory safety assessment, which is widely recognized to be in need of a transition to accommodate the increased level of toxicological knowledge and animal welfare considerations (Piersma et al., 2018b).

Received October 11, 2019; Accepted November 22, 2019;  
Epub December 11, 2019; © The Authors, 2019.

ALTEX 37(2), 208–222. doi:10.14573/altex.1910111

Correspondence: Rob Vandebriel, PhD  
Centre for Health Protection  
National Institute for Public Health and the Environment (RIVM)  
PO Box 1, 3720 BA, Bilthoven, The Netherlands  
(rob.vandebriel@rivm.nl)

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.

**Tab. 1: Overview of advantages of OC devices compared to traditional *in vitro* assays**

<b>Fluidic forces</b>	Optimal transport of soluble cues can be applied in OCs, and shear stress is exerted through medium flow. It is established that cell behavior is different in the presence of fluidic shear stress. The constant perfusion of fresh medium further maintains the pH, removes waste and provides a continuous source of nutrients.
<b>Defined concentration gradients</b>	Not only the fluid flow of the culture medium can be controlled in microfluidic devices; complex concentration gradients also can be achieved. Repeated flow stream lamination can, for example, promote cell chemotaxis, i.e., the migration of cells in response to a stimulus.
<b>Homogenous chemical distribution in medium</b>	Only a small portion of the administered dose reaches the cell membrane. Colloidal behavior, particle sedimentation, and diffusion must be taken into consideration when correcting for the initial dose at the start of the experiment. With microfluidics, a homogenous suspension of substances in culture medium can be applied to the cell membrane under continuous perfusion, thereby creating a more physiologically relevant situation.
<b>Micropatterning</b>	Micropatterning allows improved control over homo- and heterotypic cell-cell interactions and easier 3D culturing. These micropatterns can be used to control the geometry of adhesion and therefore the orientation of the cell division axis. Micropatterning is also needed for the maintenance of stem cells or the exact positioning of a cell on a sensor.
<b>Sensor integration</b>	With the integration of biosensors in the device itself, cell behavior can be monitored more reliably and quantitatively (e.g., electrical activity, cytotoxicity measurements, optical sensors, cell-based biosensors, microscale patch clamp devices).
<b>Mechanical strains</b>	Cyclic mechanical strain can accentuate toxic and inflammatory effects and enhance the transport of particles over organ barriers.
<b>Expanded cell viability</b>	Optimal control over environmental conditions for cell-based assays increases cell viability and culture time. This has been shown to allow further differentiation of the cells towards improved physiological resemblance of the tissue.

Two main approaches can be distinguished in this transition to a more predictive and animal-free safety assessment: the “evolutionary” approach and the “revolutionary” approach (Scialli et al., 2018; Burgdorf et al., 2019). The evolutionary approach represents the current practice of a more stepwise replacement of one animal test at a time by a set of non-animal methods that together predict the endpoint for which the animal test was designed. In the evolutionary approach, animal tests are often used as the gold standard. The revolutionary approach represents a more radical change in safety assessment, starting with the identification of crucial mechanisms and biochemical events in clinical toxicity observations that are relevant for humans. These mechanisms and events should be the basis for the development of a set of animal-free assays that together predict these toxicities. From this revolutionary perspective, it is conceivable that other toxicological endpoints are studied than those traditionally studied in animals. A revolution towards animal-free safety assessment is not expected to happen overnight, but gradually, parallel to the development of 3R methods that better fit into the ongoing evolutionary paradigm.

For the other part, the potential role of OCs in regulatory safety assessment depends on the opportunities they offer. To help regulators and innovators identify opportunities in the development of OCs and prioritize where resources are best used, it is crucial to identify the regulatory needs and compare the current availability, opportunities and limitations of OCs in relation to these needs. McMullen et al. (2018) have introduced a stepwise approach for this, not focusing on OCs specifically but on all *in*

*vitro* methods. While their study and ours share the same general approach, we have executed it differently, strengthening the very useful findings of McMullen et al. (2018), but also adding different insights.

The current paper gives our view on the opportunities and limitations of OCs in the context of the transition to a more predictive, animal-free regulatory safety assessment. After a brief overview of the technical promises of OCs, we discuss how OCs can fit into both an “evolutionary” as well as a “revolutionary” transition. We add to the analysis of McMullen et al. (2018) and discuss, from a regulatory viewpoint, the prioritization of organs and tissues needing an OC and/or *in vitro* assay, however, based on the analysis of a different database. In our analysis, which is based on the outcomes of a dedicated expert workshop, the regulatory needs are compared to the state of the art of OCs on several of these organs and tissues. Finally, we outline the expected challenges to implementation of this innovative technology for regulatory safety assessment purposes.

## 2 The promising features of organs-on-chip

OCs can be defined as three-dimensional (3D) tissue cultures (the mini “organ”) set in microfluidic systems (the “chip”). The microfluidic system allows flow of culture medium along or through the organ and may also allow stretching of the tissue to mimic breathing or peristaltic contraction of the intestine (Huh et al., 2010; Kim and Ingber, 2013). This flow appears to cause



cell barriers to polarize and differentiate, and to improve viability duration of cell cultures. As OCs are meant to mimic the smallest functional unit of an organ or tissue (Marx et al., 2016), multiple cell types are typically included. OCs therefore represent the most complex form of *in vitro* cell cultures; monolayers being their most simple form and 3D cultures being an intermediate form. Also named microphysiological systems, these devices have been shown to support tissue viability for 28 days or longer (e.g., Baxter et al., 2015; Maschmeyer et al., 2015; Xiao et al., 2017). They therefore overcome two important challenges in studying complex, systemic toxicity endpoints and toxicokinetics: multiple interacting cell types and longer viability. In addition, many OCs are based on human cells, which may be more relevant to address human toxicity than animal models or animal cells. OCs also allow culturing multiple tissues together, connected by medium flow (e.g., Maschmeyer et al., 2015; Oleaga et al., 2016), thus potentially allowing toxicokinetics to be addressed simultaneously with toxicity. Some additional advantages of OCs in comparison to less complex *in vitro* models are explained in Table 1.

### 3 The role of OCs in an evolutionary transition of regulatory safety assessment

The idea behind an evolutionary transition to a more predictive, animal-free regulatory safety assessment is to gradually replace all regulatory animal toxicity studies by (sets of) more predictive, animal-free methods. Currently, authorities have already accepted animal-free methods for some toxicity endpoints. For example, for assessing irritation, corrosion, sensitization, and genotoxicity potential of chemical substances, there are a number of OECD test guidelines (TG) that describe animal-free tests (Tab. 2). These OECD TGs have undergone international validation and evaluation and are accepted by all OECD member countries. The safety evaluation of pharmaceuticals makes use of technical

guidelines published by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). There is significant overlap between ICH and OECD TGs.

It is evident from Table 2 that there is a lack of accepted animal-free methods for the more complex, systemic endpoints (acute toxicity, repeated dose toxicity, carcinogenicity, reproductive toxicity and neurotoxicity). In addition, there is only one OECD TG for an *in vitro* toxicokinetics assay: OECD TG 428 for dermal absorption. To date there are no TGs for *in vitro* assays to assess absorption by inhalation or by ingestion, for liver metabolism or for kidney elimination.

To address these more complex toxicity endpoints, the human-based systems with longer viability options of OCs could potentially be of great value. Theoretically, from an evolutionary transition perspective, each animal study addressing complex toxicity endpoints could be replaced by a set of human-derived OCs of each organ evaluated in such a study. In combination with appropriate models for toxicokinetics, these OCs could then be used to derive safe limit values for humans.

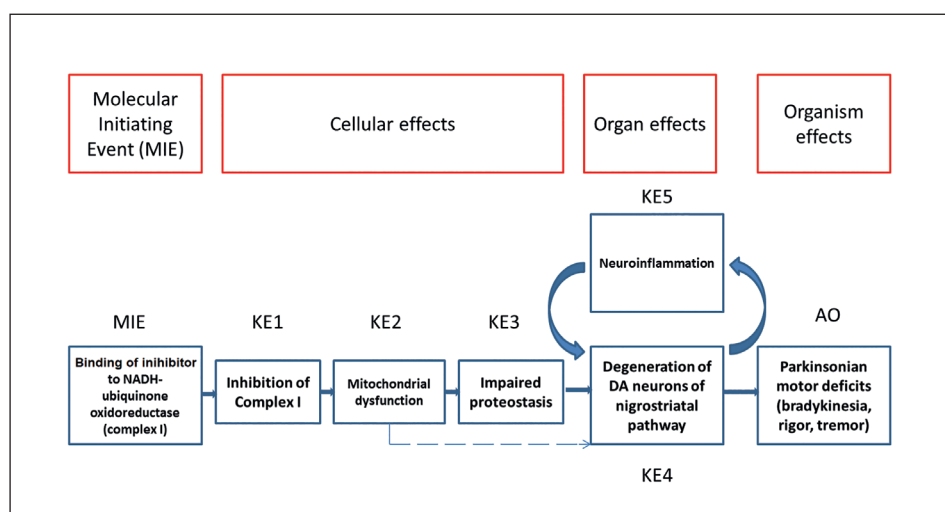
### 4 The role of OCs in a revolutionary transition of regulatory safety assessment

The revolutionary transition approach to a more predictive animal-free regulatory safety assessment does not necessarily require a one-by-one replacement of currently used animal studies. Instead, this approach can make use of the development of adverse outcome pathways (AOPs). AOPs are the basis of a mechanism-based toxicity approach. They provide a structured representation of biological events leading to adverse effects, starting with a molecular initiating event (MIE) of a chemical binding to a certain endogenous structure and leading to the adverse outcome (AO) via several key events (KEs). An example of an AOP for Parkinsonian motor deficits is shown in Figure 1.

**Tab. 2: Toxicity data generally required in regulatory dossiers of chemical substances, the OECD test guidelines for producing such data that require the use of animals, and available accepted animal-free OECD test guidelines**

Required toxicity data (by endpoint)	OECD test guideline methods using animals	Animal-free OECD test guideline methods
Acute toxicity (3 routes)	401, 402, 403, 436, 425, 423, 420, 433	
Irritation/corrosion (eye and skin)	404, 405	460, 437, 438, 491, 492, 430, 431, 435, 439
Sensitization	406, 429, 442A, 442B	442C, 442D, 442E
Repeated dose toxicity	407, 408, 409, 410, 411, 412, 413, 452	
Genotoxicity	488, 489, 483, 478, 475, 474, 473, 485, 484	471, 490, 487, 476
Carcinogenicity	451, 453	
Reproductive toxicity (fertility and developmental toxicity)	443, 414, 415, 416 (421 and 422 for screening only)	
(Neurotoxicity) <sup>a</sup>	424, 419, 418, (426)	

<sup>a</sup> Only few regulations and directives require neurotoxicity data.



**Fig. 1: The AOP “Inhibition of the mitochondrial complex I of nigrostriatal neurons leads to parkinsonian motor deficits” of Bal-Price et al. (AOP number 3 of the AOP Wiki)**

The AOP shows the molecular initiating event (MIE) of a chemical that binds to a certain endogenous structure of a human. This leads to the final adverse outcome (AO) via several key events (KEs).

Source: <https://aopwiki.org/aops/3>.

Ultimately, mechanisms of toxicity for diverse adverse health effects can be described by AOP networks, ideally in a quantitative manner (Pittman et al., 2018; Hassan et al., 2017). The toxicity of chemical and pharmaceutical substances can then be predicted by testing their effects on the MIE and on KEs using designated animal-free methods. In the example above, the ultimate AO itself (in this case Parkinsonian motor deficits) is not detectable with an *in vitro* assay, but the steps leading to this AO can be measured *in vitro* (e.g., mitochondrial dysfunction or degeneration of dopaminergic neurons of the nigrostriatal pathway).

In this revolutionary approach, OCs may have great value in identifying the MIE and KEs of an AO in humans and in studying possible links between different AOPs. In addition, OCs may be better at detecting effects of substances on MIE and KEs, especially for those events that are difficult to determine using simpler methods. As shown for the example of lung injury and disease, knowledge of the AOP underlying such AOs is necessary to know the components the model needs to include and the functions it needs to cover (McMullen et al., 2018).

## 5 Integrating OCs with other safety assessment methods

It is important to note that OCs will seldom be the only method of choice, or even the first choice from multiple available methods when it comes to replacing current animal studies. Simpler and cheaper methods may be preferred, depending on their suitability to determine the AO at hand. In most cases, a set of test methods will be needed to replace one animal study or predict an AO, as is shown by the example of skin sensitization (OECD, 2016a). Here, a set of relatively simple assays suffices to show that a substance induces the KEs involved.

The use of multiple different methods aimed to address a certain type of toxicity can be combined in an integrated testing strategy (ITS) or integrated approach to testing and assessment (IATA). In an IATA, information from multiple information

sources, such as *in chemico*, *in vitro*, *in vivo*, or omics, is integrated and converted into a prediction model. Defined approaches (DAs) can also be part of IATAs. In DAs, data are generated by a defined selection of specific animal-free methods and subsequently evaluated by means of a fixed data interpretation procedure (OECD, 2016b). For the case of skin sensitization, multiple different DAs can be designed for the same endpoint, as different assays and models can be selected and different combinations can be made. Some of the strategies for skin sensitization provide information that can be used for hazard identification, e.g., distinguishing skin sensitizers from non-sensitizers, while other approaches provide information that allows potency sub-categorization as well (Ezendam et al., 2016). The integration of an OC in an IATA may support more complex hazard assessment in cases where this is needed.

The challenge that arises is how to deal with the large number of different approaches that would need to be assessed for suitability for regulatory purposes and how to perform such validations (Piersma et al., 2018a), as will be discussed below. Recent grey literature on the subject includes two roadmaps: the US Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) strategic roadmap discusses the reasons for the paradigm shift on validation (ICCVAM, 2018) and the RIVM roadmap for animal-free innovations in regulatory safety testing describes the steps required to facilitate the transition to animal-free safety assessment including validation (RIVM, 2018).

## 6 Prioritization of organs and tissues for developing OCs

For both the evolutionary as well as the revolutionary approach, replacing the animal studies with more predictive animal-free methods will likely take decades of research. In this challenge, it may help regulators and innovators to prioritize which organs are of most interest to develop an OC. These would be the organs



and tissues that are most often involved in manifestation of AOs upon chemical or pharmaceutical exposure in humans. Again, the development of an OC should only be prioritized if simpler or more suitable methods that cover the respective AO in the target tissue or organ are not feasible. Therefore, it would be valuable to analyze which KEs in which tissues are involved in the AOs that are most relevant for humans.

Ideally, one would analyze all AOPs that predict any type of adversity in humans and subsequently identify the organs or tissues that are involved. Next, it would need to be assessed whether OCs are needed to assess effects on these KEs, or whether simpler methods exist. However, such a complete set of AOPs is not available yet, thus such an analysis is not yet possible.

Instead, as a starting point, an indication of the organs most involved in determining adversity by chemical and pharmaceutical substances can be retrieved by studying data of animal toxicity studies. McMullen et al. (2018) performed such an analysis on a collection of databases in the USA (the “database collection”), studying which systems most often led to a regulatory decision. They found that the liver was involved in most regulatory decisions, followed by the neurological, urinary and respiratory systems. While this provides a very useful insight, a limitation of this approach is that it neglected which other organs may have shown an effect, even though it was not at a level leading to a regulatory decision. We take a similar, but slightly different approach, based on an analysis of the RepDose database by Batke et al. (2013) and considering which additions are necessary for complex endpoints other than repeated dose toxicity. This approach is discussed below.

#### *Repeated dose toxicity*

The current OECD guideline for repeated dose toxicity requires the evaluation of 21 tissues, but it is questionable whether analysis of all these tissues is needed. A study systematically analyzing a database of 88 oral repeated dose toxicity studies on chemical substances showed that for 86% of the test substances, the lowest observed adverse effect level (LOAEL) can be derived by analyzing only six organs/whole body symptoms (Batke et al., 2013). A ranking of the organs/whole body symptoms based on how often these showed an effect at the dose identified as the LOAEL is given in Table 3. The top 10 organs most involved in determining adverse effect levels in oral repeated dose toxicity studies are marked in bold. If simpler methods are not taken into account, these 10 organs offer a starting point to prioritize for OC development, with the following further considerations:

- The forestomach is excluded from this top 10, as humans do not have this organ. It is nevertheless analyzed in animal studies because it is the site of first contact for substances given by gavage, thus providing information on first-contact effects.
- The dataset used by Batke et al. (2013) did not include inhalation and dermal studies, where most probably the lung and skin, respectively, would have ended up high in this ranking. Thus, these two organs were chosen to be added at the top of this table. It is noted that inhalation and dermal studies might have given a different order in the ranking of the other organs, had they been included.

- As the OECD TG for the 90-d oral study was revised in 1998 to derive more information from the study, it is possible that older studies in the RepDose database do not include as many parameters as newer studies, which would also affect the ranking. Furthermore, additional toxicity endpoints, such as endocrine disruption, currently are being considered for inclusion in regulatory testing. This may also shift the priority for OC development to other or additional tissues.
- These are data from rats, and the ranking might be different for humans.
- Body weight change and clinical symptoms cannot be analyzed using *in vitro* methods such as OCs. Often, these symptoms will be the indirect result of effects in tissues and organs, in which case they are expected to be covered by analyzing relevant biomarkers in these tissues and organs. In some cases, symptoms such as a drop in body weight or lethargy may be the first sign of distress, before any changes in currently used clinical biomarkers can be measured. The development of more sensitive parameters, such as early KEs, may overcome this limitation.

The oral studies in the RepDose database showed that liver was most often involved in the manifestation of AOs, followed by kidney and testes. By assumption, lung and skin would also rank highly if inhalation and dermal studies had been included. This may be in line with the findings of McMullen et al. (2018), in whose database collection dermal studies are likely rare, possibly explaining that skin does not rank highly (McMullen et al., 2018). The difference in ranking between the testes and the neurological system is interesting and can as yet not be explained. The fact that only repeated dose toxicity studies were included in our study, while all types of toxicity studies were included in the analysis of McMullen et al. (2018), does not explain this difference. To our knowledge, no similar analysis to that of Batke et al. (2013) of organs and tissues most often involved in manifestation of AOs exists for other complex, systemic endpoints such as reproductive toxicity and immunotoxicity. The organs and tissues most involved in these endpoints may be different compared to those involved in repeated dose toxicity, as will be discussed below.

#### *Reproductive toxicity*

Reproductive toxicity involves effects on both fertility and developmental toxicity. Some effects on fertility, such as testes atrophy, can already be picked up by analysis of the gonads and uterus in a repeated dose testing setup. However, for a more comprehensive assessment of effects on fertility, all organs involved in fertility would need to be covered by OCs and, most importantly, they would need to be interconnected as a delicate hormone balance steers these processes. At the very least, an OC for the uterus, gonads (male and female) and fallopian tube would be necessary. One would expect that the organs involved in the reproductive endocrine system are necessary too, such as the hypothalamus and anterior pituitary. A full menstrual cycle with oocyte maturation in a microfluidic multi-organ setup without pituitary and hypothalamus has been reported (Xiao et al., 2017), however this development remains to be confirmed by further study. In any case, for fertility testing, single organs or tissues

**Tab. 3: Ranking of targets in 90-d oral rat studies listed according to how often they showed an effect at the LOAEL**

Adapted from Batke et al. (2013). The availability of proof-of-concept OCs is indicated for the top ten organs and tissues (bold). The availability of OC was concluded by experts at a dedicated workshop at RIVM in 2017 (see below). In the meantime, an OC model has been developed for stomach (Lee et al., 2018).

Target (whole body symptom/organ effect/blood marker effect)	% of studies where the target was affected at the LOAEL	OC developed as proof-of-principle
<b>Lung</b>	n.d.	Yes (e.g., Huh et al., 2010)
<b>Skin</b>	n.d.	Yes (e.g., O'Neill et al., 2008)
Clinical chemistry	38.6	
<b>Liver</b>	31.8	Yes (e.g., Wagner et al., 2013)
Haematology	29.5	
Kidney	23.9	Yes (e.g., Vedula et al., 2017)
Body weight	21.6	
Clinical symptoms	18.2	
<b>Testes</b>	8.0	No
Urine analysis	6.8	
Forestomach	6.8	
<b>Heart</b>	5.7	Yes (e.g., Grosberg et al., 2011)
<b>Spleen</b>	4.5	No
<b>Brain</b>	2.3	No
<b>Stomach</b>	2.3	No
<b>Intestine</b>	2.3	Yes (e.g., Kim and Ingber, 2013)
Adrenal gland	1.1	
Thyroid gland	1.1	
Thymus	1.1	
Lymph node	1.1	
Bladder	1.1	
Endocrine system	1.1	
Nose	1.1	
Pancreas	1.1	
Vascular system	1.1	
Salivary gland	1.1	
Pituitary gland	0	
Nervous system	0	
Ovary	0	
Bone	0	
Gallbladder/bile duct	0	
Uterus	0	
Prostate	0	
Seminal vesicle	0	
Skeletal muscle	0	
Adipose tissue	0	
Eye	0	



will not suffice; multi-organ setups with appropriate connections between the tissues are necessary.

Achieving an adequate OC for developmental toxicity testing may not be feasible at all, as this would imply the culture and exposure of human embryos past the initial stages, which runs into ethical barriers. Other solutions are probably more appropriate, such as zebrafish embryo testing or computational modelling.

#### *Carcinogenicity*

Carcinogenicity already can be detected in part by animal-free genotoxicity assays, although in the EU, *in vitro* testing alone is currently not accepted to be sufficient to address mutagenicity or genotoxic carcinogenicity, except for cosmetics. No animal-free method is currently available for the detection of non-genotoxic carcinogens. Non-genotoxic carcinogenicity may arise in any tissue, with some tissues having a higher exposure and thus a higher risk, or with some tissues being more sensitive to the carcinogenic effects of certain types of substances. For example, estrogens can cause cancer mostly in tissues with high levels of a certain type of estrogen receptor, such as breast tissue. Non-genotoxic carcinogenic mechanisms can be picked up through transcriptomic, biochemical, and cell phenotypic analyses, which can be performed in the same *in vitro* tests as those used for repeated dose testing. As these changes occur much earlier than the emergence of a tumor, it is not to be expected that the test duration would need to be longer than that for repeated dose testing, thus merging these two endpoints. Whether additional tissues/organs would need to be included remains to be determined.

#### *Neurotoxicity and immunotoxicity*

Organs and tissues involved in neurotoxicity and immunotoxicity are included to some extent in current repeated dose testing, but this is mostly limited to histopathological evaluation. Due to increasing concerns about these endpoints, one could argue that more functional assays are needed. For neurotoxicity, obviously brain tissue and nerves are relevant, which have already been identified as important tissues involved in repeated dose toxicity. For immunotoxicity, spleen, thymus, lymph nodes, and bone marrow are important tissues, some of which were also already identified as important in determining repeated dose toxicity. However, the immune system is much more complex, and proper functioning is dependent on interaction between the different tissues. For this endpoint, determining the necessary tissues and their interrelations therefore deserves a more careful analysis.

#### *Cardiotoxicity*

While the heart is included in repeated dose testing of chemical substances, this is mostly limited to histopathological evaluation. For pharmaceuticals, cardiotoxicity testing is much more elaborate, including functional testing, as cardiotoxicity is one of the main reasons for product failure at relatively late stages of pharmaceutical drug development. However, the current testing strategy, which is in part based on animal testing, is a subject of debate because of its poor predictive value (Pridgeon et al., 2018). This in itself is a good reason to develop better models such as a heart-on-a-chip. For non-pharmaceuticals, the relevance of car-

diotoxicity for humans is unknown due to a lack of regulatory testing requirements.

#### *Toxicokinetics endpoints*

The major toxicokinetics endpoints for which animal-free methods are necessary are lung absorption, skin absorption, intestinal absorption, liver metabolism, protein binding, blood-tissue partitioning, and translocation over internal barriers, such as the blood-brain barrier, blood-testis barrier, and placenta. As stated earlier, an OECD TG already exists for skin absorption. Partitioning and protein binding (together dictating distribution) can be determined from the octanol-water partition coefficient (K<sub>ow</sub>) and a serum protein binding assay that does not need tissue culture. Internal barrier passage, such as the blood-brain barrier, blood-testis barrier, and placenta, are mostly relevant for the endpoints neurotoxicity, fertility, and developmental toxicity, respectively. Some organs and tissues involved in kinetics, such as lung, liver, and kidney, have already been identified as a priority because of their important involvement in repeated dose toxicity. It should be noted that the OCs to address repeated dose toxicity in these tissues and organs may not be suitable to address toxicokinetics. For example, for the evaluation of absorption, two compartments, one on each side of the tissue, are required, while for repeated dose toxicity, only one compartment would suffice. More refined toxicokinetics assessment can be performed when parameters such as bile excretion and sweat excretion are known, but these kinetic endpoints are of lower priority as they influence the total toxicokinetics to a lesser extent than the other parameters mentioned (Bessems et al., 2014).

## **7 Workshop comparing regulatory toxicity testing needs to current state of the art of OCs**

In June 2017, a workshop was held at RIVM to assess to what extent the current state-of-the-art of OC development met the needs of regulatory toxicity testing. The workshop was focused on ten tissues, organs, and organ systems that are amongst those of highest importance for current regulatory toxicity testing (based on the above analysis): lung, skin, liver, kidney, heart, intestine, reproductive system, immune system, neuronal system, and vascular system. Regulators and OC developers together discussed the needs, opportunities and limitations of developing OCs for these ten tissues.

#### *State of the art*

The workshop brought forward that OCs for lung, skin, liver, kidney, heart and intestine are already available (e.g., Grosberg et al., 2011; Huh et al., 2010; Kim and Ingber, 2013; O'Neill et al., 2008; Vedula et al., 2017), even combined in multi-organ chips (e.g., Maschmeyer et al., 2015; Oleaga et al., 2016; Wagner et al., 2013) (Tab. 3). Together, these organs and tissues may already be able to cover aspects of kinetics, cardiotoxicity and repeated dose toxicity. The workshop participants agreed that they had reached a stage of development that required input from regulators on the precise needs in terms of features and functionalities per organ.

**Tab. 4: Recommended priority list of research and development issues per organ or functional system**

The order of the organ or functional system is random and does not follow any prioritization.

Organ	Priority issues to be tackled
Lung	<ul style="list-style-type: none"> <li>– Barrier function for kinetics, especially in alveoli</li> <li>– Better alveolar models</li> <li>– Mucociliary clearance</li> </ul>
Skin	<ul style="list-style-type: none"> <li>– Dermal absorption</li> <li>– Potency</li> <li>– Vasculature for 2-way migration of immune cells, introduce defined immune cells into the skin</li> </ul>
Liver	<ul style="list-style-type: none"> <li>– Connection to endothelial cells and generation of bile ducts</li> <li>– Consensus set of biomarkers for liver toxicity</li> </ul>
Kidney	<ul style="list-style-type: none"> <li>– Combining of glomerulus and proximal tubuli</li> <li>– Consensus set of biomarkers for kidney toxicity</li> </ul>
Heart	<ul style="list-style-type: none"> <li>– Cardiac action potential analysis in the chip</li> <li>– Validation versus <i>in vivo</i> and versus simpler systems</li> </ul>
Intestine	<ul style="list-style-type: none"> <li>– Inclusion of microbiome</li> <li>– Barrier function for kinetics</li> <li>– 3D models</li> </ul>
Testes and other parts of the reproductive system	<ul style="list-style-type: none"> <li>– Mimic one reproductive cycle, incl. oocyte maturation, still to be done for spermatogenesis<sup>a</sup></li> <li>– Mimic blood barrier function</li> <li>– Sperm function</li> </ul>
Spleen, thymus and other parts of the immune system	<ul style="list-style-type: none"> <li>– Determine immune function of the various organs, to put function-on-a-chip rather than organ-on-a-chip</li> <li>– Multiple organs-on-a-chip mimicking immune response</li> <li>– Germinal center-on-a-chip</li> <li>– <i>In vitro</i> adaptive immune response</li> </ul>
Brain and other parts of the neuronal system	<ul style="list-style-type: none"> <li>– If OC needed, high volume system (analytical issue with detection limits)</li> <li>– For 3D brain, relation readouts with neurocognitive effects<sup>b</sup></li> </ul>
Vascular system	<ul style="list-style-type: none"> <li>– Readouts for plaque formation and inflammation</li> <li>– Contraction and dilatation analysis</li> <li>– Control of pressure of vessel barrier and flow control through lumen</li> <li>– Angiogenesis</li> </ul>

<sup>a</sup> Already done for oocyte maturation (Xiao et al., 2017); <sup>b</sup> A more complex system (layered structure) is needed for a higher cognitive function.

For the other organ systems (reproductive system, immune system, neuronal system and vascular system), further research and development would be necessary first.

For the reproductive system, the OC state-of-the-art is a multiple organ system that has been reported to mimic a menstrual cycle (Xiao et al., 2017). Although a major achievement, it still comprises only a small part of the whole reproductive process. The possibilities for mimicking the whole reproductive process are severely hampered by the difficulties in culturing the gonads, and the complex nature of the reproductive system, which necessitates multiple connected tissues and establishment of a delicate hormone balance. While technical challenges of culturing gonads and connecting multiple tissues may be solved with microfluidic systems, there are ethical restraints in culturing human stem cells and embryos, and thus limitations to what can be achieved with regard to mimicking the reproductive system using *in vitro* assays such as OCs.

Immunotoxicity is a collective name for many types of adverse immune effects, such as hypersensitivity, inflammation, immunosuppression, and auto-immunity. Some relatively simple *in vitro* assays are already available that address certain types of immunotoxicity, while other types are much harder to address with simple *in vitro* assays because, for example, multiple organs and tissues are involved. The workshop participants recommended that a more thorough analysis is needed to identify the functionalities in immunotoxicity, after which it can be determined which organs and tissues are necessary to cover these.

In order to address neurotoxicity, a potential OC would need to be able to process information and transduce signals in ways similar to the human brain. However, the brain with its associated neural system is of such high complexity that it is understood relatively poorly, making the development of a system that can mimic it in the form of an OC extremely challenging. It was also questioned whether the level of an OC would be necessary in



this case, since a 3D culture format without microfluidics might suffice.

With regard to the vascular system, various research groups have already developed functional hearts-on-a-chip for various purposes. Others have focused on developing blood vessels to be integrated into organ cultures. Nevertheless, systems that are able to mimic blood pressure changes, vascular dilation and contraction, plaque formation and angiogenesis are not yet available.

#### Priorities in further research

At the RIVM workshop, small discussion groups of regulators and developers held a further in-depth exchange on the features and functionalities that are necessary (from the regulators' perspective) and are feasible (from the developers' perspective) for each of the ten organs or functional systems selected. These issues, presented per organ or functional system in Table 4, can be considered priorities in further research.

#### Summary

The insights obtained in this workshop led to recommendations for regulators on where their involvement is required when OCs are past the research and development stage and ready to move towards optimization and possible implementation. The discussions showed that, currently, regulators may play an important role in the optimization of the lung, intestine, liver, and kidney for toxicokinetics evaluation and repeated dose toxicity testing, as well as in the optimization of the heart for cardiotoxicity (= repeated dose toxicity) testing. Here, aspects such as the necessary

level of complexity and the ease of application should be considered, after which further steps of standardization, reproducibility testing, etc. can be undertaken towards acceptance and regulatory implementation. Furthermore, input of the workshop participants led to a list of priority issues per organ or functional system (Tab. 4) to be tackled in further research and development of OCs.

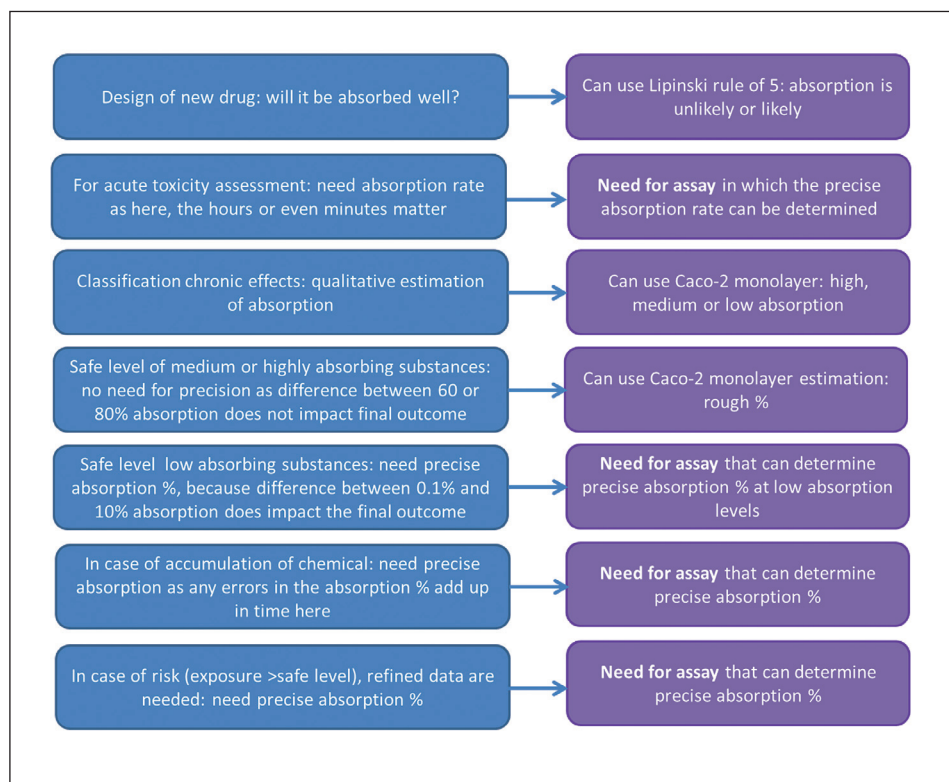
This outcome is somewhat different from that of McMullen et al. (2018), who found that a key shortcoming in the current efforts is the ability to test volatile compounds and to predict pulmonary toxicity and therefore advise to focus on an *in vitro* system for the lung. While this is part of what we have found, too, our analysis was somewhat broader, e.g., also considering the needs for toxicokinetics, and we have differentiated between which OCs are at a stage at which regulators should become involved in their optimization and which OCs first need further (academic) development.

## 8 Technical recommendations for regulatory OCs

Apart from prioritizing the type of organs and tissues, a number of technical recommendations can be given to facilitate the development of OCs that can be successfully implemented in a regulatory context.

#### Predictive value

An OC needs to be able to provide predictive information on the hazard or toxicokinetics of a chemical or pharmaceutical sub-



**Fig. 2: Different information needs and related possibilities for oral absorption assays**

stance. The level of accuracy of this prediction depends in part on the type of regulatory safety assessment question at hand.

To illustrate this, different questions and related predictive information requirements are shown for oral absorption of a chemical or pharmaceutical substance (Fig. 2). This example shows how for certain regulatory questions, there is a need for an assay that can determine the precise rate and percentage of absorption of chemicals, while for others, a rough estimate will suffice. It is likely that OCs will have more value compared to simpler *in vitro* systems as the level of detail in the information required increases.

#### *Level of complexity*

Closely related to the predictive value is the required level of complexity of an OC. One option (whether realistic or not) is to make an OC resemble human physiology as closely as possible, based on the hallmark of biology that structure and function are interrelated. A one-on-one resemblance to real tissue implies proper coverage of human physiology and KEs that are involved in AOPs. Achieving such close resemblance to real tissue would involve, e.g., including multiple cell types and blood vessels in a 3D architecture on bio-based scaffolding with peristaltic movements in a flow-through system. However, this level of complexity is not always needed and even may have disadvantages such as higher costs, longer production times, higher failure rates, and lower reproducibility between experiments and between labs. Microfluidics, for instance, may not be necessary for all purposes in regulatory safety assessment. Likewise, it may seem advantageous to measure as many parameters as possible in these systems, but the generation of large amounts of data may not be valuable or may even be counterproductive if the data does not provide clear, unambiguous answers. More data may sometimes increase the uncertainty in the outcome of a test if the biological relevance of the data is not prioritized appropriately.

Therefore, it is recommended to find the right balance between complexity (high biological accuracy, multi-organ-on-a-chip to human-on-a-chip) and ease, resulting in an OC that is fit-for-purpose for regulatory needs. Again, this balance depends on the regulatory question at hand, as discussed above. Thus, multiple systems with differing levels of complexity may need to be available. For illustration purposes, some examples of different levels of complexity are discussed in Box 1 (absorption, liver clearance rate, liver fibrosis). In short, the advice is: Keep *in vitro* assays as simple as possible, but as complex as necessary.

#### **Box 1:** **Absorption**

For oral absorption, it has been found that the 2D monolayers of Caco-2 cells provide some qualitative information on the level of absorption of substances, but they are not suitable for precise determinations (Fig. 2). Addition of mucus-producing cells or flow of the medium has been

shown to impact the absorption of chemicals and nanoparticles (Shim et al., 2017). Addition of a 3D collagen scaffold on which the Caco-2 cells are grown leads to the production of mucus (Kim et al., 2014) and changes absorption of chemicals, both in static culture (Yu et al., 2012) and in cultures with flow (Shim et al., 2017). Medium flow plus cyclic mechanical strain mimicking peristaltic movement was shown to induce the formation of villi-like structures in Caco-2 cells and to lead to differentiation into different cell types, including mucus-producing Goblet cells (Kim and Ingber, 2013), which impact absorption. The expression of metabolic enzymes in the cells was also changed (Kim and Ingber, 2013; Shim et al., 2017), which affected the apparent absorption of chemicals. For microparticles, the uptake may occur mainly through M-cells in the Peyer's patches in the intestine (Brun et al., 2014), which would therefore need to be present to accurately determine the absorption of such particles. Finally, the presence of gut microbiota may also affect the (apparent) absorption. Thus, there are multiple factors that can be used to increase the complexity of the gut culture, potentially leading to a more accurate determination of absorption. It remains to be investigated which of these factors have the most impact and thus which minimal additions of complexity can lead to a sufficiently accurate and precise determination of absorption.

#### **Liver clearance rate**

Liver clearance has been found to be reasonably well determined with suspensions of primary hepatocytes or HepaRG cells in case of fast to medium-fast metabolizing substances (Gouliarmou et al., 2018). For slowly metabolizing substances, these cultures do not stay functional long enough (Lauschke et al., 2016). There is, therefore, a need for a culture system that allows longer viability of hepatic cells while maintaining metabolic enzyme activity. It has been shown that liver spheroids retain morphology, viability, and hepatocyte-specific functions in a static culture for culture periods of at least 5 weeks (Bell et al., 2016). It now remains to be verified whether the clearance of slowly metabolizing substances can be predicted with sufficient accuracy with such a system or whether additional complexity is necessary, e.g., by addition of other liver cell types or addition of medium flow.

#### **Liver fibrosis**

For mimicking the process of liver fibrosis, as one of the main types of adverse effects in the liver, it has been shown that a static co-culture of three liver cell types is sufficient (Norona et al., 2016). This bioprinted liver tissue, made with hepatocytes, stellate cells, and endothelial cells, showed deposition of collagen after 14 days of treatment with fibrogenic agents. The flow of an OC therefore does not seem to be necessary. However, a simple culture of hepatocytes alone is not complex enough to obtain fibrosis.



### Longevity

It was suggested that toxicity testing in OC models should align to the exposure periods dictated for animal testing (acute, sub-acute (28-d), subchronic (90-d), possibly even chronic) (Mastrangeli et al., 2019). However, this may not be necessary when OCs make use of readouts that allow for effects to be picked up at earlier time points, rather than mimicking the apical effects that are currently determined in animal studies. Here, the importance of developing AOPs is underlined, as evaluation of effects of substances on MIEs and KEs may facilitate the early detection of toxicity.

On the other hand, one should be aware that the mechanisms ensuring homeostasis are more comprehensive, diverse, and layered in a complete organism compared to an OC, even when the OC consists of multiple interacting cell types. The different events leading up to an AO may occur at different time points, i.e., to detect all KEs involved in an AOP in a single OC, long exposure durations may still be necessary. Therefore, the required longevity of a particular OC should be established on a case-by-case basis, may differ between different OCs, and may well not be the same as for animal studies.

## 9 Implementation challenges

Apart from the technical challenges and recommendations discussed above, several challenges exist in implementing these complex systems in regulatory toxicity testing, with regard to their standardization, validation, cost, and other issues.

### Standardization

After optimization of an OC to make it fit-for-purpose, as described above, it needs to be standardized. Standardization of an OC device involves the process of achieving uniform cells or tissue types that are compatible with chips and preferably can form compatible modules for interconnection into a multi-organ system. Moreover, standardization involves describing the preparation and use of an OC in such detail in an SOP that the results of the readouts are robust and reproducible. For complex systems such as an OC, a rigorous and strict standardization of the methods will be necessary to achieve sufficient reproducibility. Helpful guidance documents for this are the OECD Guidance Documents (GD) 211 (OECD, 2014) and Good In Vitro Method Practices (OECD, 2018) on how to describe an *in vitro* assay and on what aspects and possible artefacts to consider when developing an *in vitro* method, respectively.

### Validation

Some form of validation process will be needed for new test methods such as OCs before they can be implemented into a regulatory safety assessment context. Generally, this validation involves a long process of standardization and harmonization of the method, eventually leading to inclusion into technical guidelines such as those from the OECD or ICH, which pave the way for integration into legal frameworks by way of the mutual acceptance of data (MAD) principle. For example, accep-

tance within the OECD TG program requires an assessment of the relevance and reliability of a method, which is described in OECD GD 34. One aspect in this classical validation approach is the comparison of the results of the new method with those of the “old” method, the gold standard, by testing a variety of reference chemicals.

This poses the question of how this validation process can be achieved for new technologies such as complex OCs and other methods and approaches aimed to replace animal studies. Most of these methods only cover part of a toxicity endpoint, so the results of the new method will not fully match those of the old method.

For this reason, the current validation process as a whole is now under debate. It is proposed that, in case of IATAs, quality criteria of reproducibility and transferability, and description of the chemical applicability domain remain essential at the level of individual assays (Piersma et al., 2018a). However, validation in terms of predictive value of individual assays, based on a variety of chemicals, is no longer opportune. Rather, one could validate that the approach (e.g., IATA) offers sufficient coverage of the biological domain, i.e., by ensuring that all pathways that are relevant to, e.g., liver toxicity, are included in the new approach with proven functionality, using a set of assays. A challenge is that multiple IATAs may be developed for the same toxicity endpoint, which would necessitate multiple validations. Clearly, this issue still needs further deliberation.

On another note, a more practical barrier for validation of OCs seems to be that in many cases the ownership of the complete OC system lies with many different parties. Different companies develop and market the different components of an OC system: the microfluidic chip, the cells/tissue, and any scaffold (including gels) at minimum, and sometimes also more specialized components, e.g., for sensing. None is thus willing to invest alone in the validation of the complete OC, as the risk would then fall to one company, while others profit without risk. Multi-partner projects with a consortium of companies, each required to invest in part of the OC, could be a solution to breach this deadlock, perhaps under direction and with subsidy of the European Commission in order to support the European business in this field. Involvement of regulators could help steer the validation towards regulatory acceptance, although it must be safeguarded that regulators do not give advantage to individual companies (see below).

### Cost

Discussing the specific cost of OCs is difficult at this stage of development, but some general aspects may already be considered. The cost of OCs themselves will in part depend on whether they are purchased as plates/chips (requiring much in-house labor to add cells or tissues), or as ready-to-use models. Compared to more conventional, simple *in vitro* assays, OCs sometimes require expensive media and possibly expensive scaffolds for their culture. They may also require longer times to prepare, and often make use of longer test durations than other *in vitro* assays. Equipment for readouts is generally similar to that needed for other *in vitro* assays, thus not incurring much additional cost. However, overall, the use of OCs will probably cost more than more conventional, simple *in vitro* assays.

On the other hand, the overall cost of OC may be lower when compared to animal tests, as the breeding and housing costs of animals may be higher and experimental durations may be longer. The cost of analyzing the different parameters in multiple tissues is not expected to be very different between OC systems and the current animal tests. However, the comparison of costs between OCs and animal testing strongly depends on the set(s) of OCs needed to replace animal tests.

#### *Intellectual property*

The competitiveness of many OC start-ups relies largely on the intellectual property (IP) licensed or owned. Since most OC devices can be easily reverse engineered, these are often not useful to protect by IP. Rather, tissue compositions, culture media compositions, hardware (such as sensors), and software can be protected by IP (Zhang and Radisic, 2017). Since patent applications are not disclosed for a long time period, it is difficult to obtain updates on the current developments of a company.

Switching from polydimethylsiloxane (PDMS) prototyping (as PDMS is not a suitable material because it absorbs many chemicals) and small-series to large-scale production of other materials requires redesign steps that will be very expensive for the relatively small companies involved. OC companies therefore need external partners and high financial investments to pass through the prototyping phase and to scale up production. Private investors seem to have gained confidence in OC technology; in Europe, at least two companies have succeeded to raise investor funds: Mimetas (\$27.65M) and TissUse (\$4.6M).

A different issue related to IP is the fact that the OECD is reluctant to accept guidelines that include models that are covered by IP, because they do not want to give any industry a monopoly position.

#### *Availability of cells*

Most OCs that are currently available use human cells because they are expected to better predict human physiology than animal cells. If eventually OCs are implemented in a regulatory safety assessment context, this will significantly increase the demand for human cells. Furthermore, the use of primary cells is preferred because immortalized cell lines may deviate from the primary cells they originate from to such an extent that they lose their predictive potential. Researchers are now putting a lot of effort into obtaining differentiated cultures from human stem cells, as these can provide an unlimited supply of “natural” human cells. However, stem cell-derived pure, fully matured, differentiated cells have not yet been obtained *in vitro* for all cell types. Often these cultures also contain immature cells, which may not react in the same way as mature cells. This problem may be solved within a decade, judged by the speed of developments in this field.

In the meantime, it is worthwhile to verify whether sufficient primary human cells may be available to enable replacement of animal testing. Enquiries among Dutch researchers learned that there are several biobanks, which can provide frozen human tissues but hardly any fresh tissues (Fentener van Vlissingen et al., 2017). Furthermore, there is an abundant potential supply of vi-

tal, non-frozen human tissue available at academic and peripheral hospitals of tissues such as skin, fat, blood (plasma) and uteri. Fresh colon, liver and lung tissue, originating, e.g., from tumor resections, are less commonly available.

Barriers to the ability to use fresh human tissue were mostly stated to lie in organizational and logistic problems. For example, in the Netherlands, there is no national standardized procedure to inform patients and obtain informed consent for further use of their tissue in research. In addition, hardly any (inter-)national or regional supply chains have been established. The European Network of Research Tissue Banks (ENRTB) is a first step in this direction, but further steps are necessary. The system that enables the availability of fresh, viable human tissue to researchers currently works via personal connections and interests of doctors, pathologists, and scientists, and is governed by standard operating procedures (SOPs) specifically designed for each academic hospital.

In summary, for many primary human cells, their availability for OCs is more an issue of organizing a coordinated supply chain and resolving informed consent issues rather than the actual quantity of cells available. For other tissues, such as colon, liver, and lung, availability may be more of an issue. In these cases, enabling the consent to use donated transplantation organs for toxicity testing if they cannot be used for transplantation, may be a potential solution. An effort from governmental side to work on this consent, with consideration of the ethical issues, is advised.

#### *Ethical issues with the use of OCs*

Currently, the use and sacrifice of animals for toxicity testing of chemical substances and pharmaceuticals is common practice, and is even required by many legislative frameworks. It is, however, by no means ethically neutral. Similar to the development of organoids, OCs hold the promise of a future without animal testing or at least to contribute to the replacement, reduction and refinement of animal experiments for regulatory toxicity testing (Bredenoord et al., 2017). Therefore, from the viewpoint of animal welfare, one may consider the use of OCs for toxicity testing more ethical than the use of animals. On the other hand, OCs may be associated with a different set of ethical issues.

Firstly, there is an ongoing debate on whether it is justified to create and use human embryonic tissue for research (Bredenoord et al., 2017), which is also relevant for OCs in case they make use of human embryonic stem cells. Positions in this debate vary from a human embryo deserving protection starting from fertilization, to the position of allowing the use of embryos for research under the condition that the providers of the gametes have given consent. Most countries allow embryo research under strict conditions, but do not allow deliberate creation of human embryos solely for research purposes. Instead, they restrict the use of embryos to those that are left over from *in vitro* fertilization procedures.

Secondly, when multi-organ-on-a-chip devices are further developed, they may reach a point where they mimic a human to such an extent that the question may arise whether such systems should be seen as a human life form just as embryos are from a certain developmental stage. In such a case, testing using these



systems may be considered even more unethical than testing on animals.

Lastly, the use of human (stem) cells, whether derived from embryos or adults, may raise issues of ownership and use, and the need for informed consent by the donor of the cells. As discussed above, this could affect the availability of cells.

## 10 Conclusions and outlook

OCs may play a role in the future regulatory safety assessment of chemical and pharmaceutical substances, especially in the assessment of complex, systemic toxicity. OCs offer novel possibilities, such as longer exposure durations and inclusion of multiple cell types, which might eventually allow the prediction of complex, systemic AOs for which no non-animal approaches currently exist. In addition, OCs may be of value to study the toxicokinetics of a chemical or pharmaceutical substance, e.g., by mimicking biological barriers and systems of metabolism. Clearly, there is still a lot of development work needed before OCs can be used for toxicological safety assessment. We recommend that the development of OCs goes hand-in-hand with the development of AOPs, and that they are combined with other safety information into testing strategies such as IATAs.

From a preliminary example analysis of a repeated dose toxicity database, ten organs of priority for OC development for regulatory use were identified. For a number of these organs (lung, skin, liver, kidney, heart and intestine), OCs are already at rather advanced stages of development, such that involvement of regulators is valuable in the optimization of these methods towards fitness-for-purpose. Regulatory risk assessors need to provide input on how the information generated from these systems can be integrated into a safety assessment, e.g., how the data can be used to derive a safe exposure limit or a classification and labelling category.

For organs such as spleen and testis, OCs are much more premature, if existing at all. Therefore, work on these organs will be mostly in the academic arena for the coming time. A major challenge is that for various toxicity endpoints, OCs of multiple organs need to be connected in some way to allow for interaction via metabolites, hormones, neurotransmitters, cytokines, and other molecules regulating biological processes between organs.

A number of technical recommendations may be given: OCs (or in fact any *in vitro* method) should not become overly complex but rather fit-for-purpose. For this, it must become clear what the regulatory purpose is exactly. The necessary duration of OC studies will probably be shorter than for current complex systemic animal studies. The implementation of OCs will face several challenges that need to be addressed: the current validation process does not seem opportune for OCs; there is a need for capital investment to reach final implementation, while each company only delivers part of an OC and therefore is not motivated to invest; there are some ethical issues; and the availability of human cells may become limiting.

In conclusion, while there is still much to be done, we feel that the development of OCs towards becoming part of regula-

tory safety assessment can benefit significantly from regular interactions between all stakeholders involved (see also Bos et al., 2020). In this respect, workshops attended by developers, toxicologists and regulators, where information is exchanged on how OCs can be used for regulatory purposes, have proven to be a useful format for such interaction.

## References

- Adler, S., Basketter, D., Creton, S. et al. (2011). Alternative (non-animal) methods for cosmetics testing: Current status and future prospects – 2010. *Arch Toxicol* 85, 367-485. doi:10.1007/s00204-011-0693-2
- Arrowsmith, J. and Miller, P. (2013). Trial watch: Phase II and phase III attrition rates 2011-2012. *Nat Rev Drug Discov* 12, 569. doi:10.1038/nrd4090
- Batke, M., Aldenberg, T., Escher, S. et al. (2013). Relevance of non-guideline studies for risk assessment: The coverage model based on most frequent targets in repeated dose toxicity studies. *Toxicol Lett* 218, 293-298. doi:10.1016/j.toxlet.2012.09.002
- Baxter, A., Thain, S., Banerjee, A. et al. (2015). Targeted omics analyses, and metabolic enzyme activity assays demonstrate maintenance of key mucociliary characteristics in long term cultures of reconstituted human airway epithelia. *Toxicol In Vitro* 29, 864-875. doi:10.1016/j.tiv.2015.03.004
- Bell, C. C., Hendriks, D. F., Moro, S. M. et al. (2016). Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease. *Sci Rep* 6, 25187. doi:10.1038/srep25187
- Bessemers, J. G., Loizou, G., Krishnan, K. et al. (2014). PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment: Recommendations from a joint EPAA – EURL ECVAM ADME workshop. *Regul Toxicol Pharmacol* 68, 119-139. doi:10.1016/j.yrtph.2013.11.008
- Bos, P., Geraets, L., de Wit-Bos, L. et al. (2020). Towards an animal-free human health assessment: Starting from the current regulatory needs. *ALTEX*, Epub ahead of print. doi:10.14573/altex.1912041
- Bredenoord, A. L., Clevers, H. and Knoblich, J. A. (2017). Human tissues in a dish: The research and ethical implications of organoid technology. *Science* 355, eaaf9414. doi:10.1126/science.aaf9414
- Brun, E., Barreau, F., Veronesi, G. et al. (2014). Titanium dioxide nanoparticle impact and translocation through ex vivo, in vivo and in vitro gut epithelia. *Part Fibre Toxicol* 11, 13. doi:10.1186/1743-8977-11-13
- Burgdorf, T., Piersma, A. H., Landsiedel, R. et al. (2019). Workshop on the validation and regulatory acceptance of innovative 3R approaches in regulatory toxicology – Evolution versus revolution. *Toxicol In Vitro* 59, 1-11. doi:10.1016/j.tiv.2019.03.039
- EC – European Commission (2013). Report from the Commission to the Council and the European Parliament: Seventh Report on the Statistics on the Number of Animals Used for Experimental and Other Scientific Purposes in the Mem-

- ber States of the European Union. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013DC0859&from=EN>
- Ezendam, J., Braakhuis, H. M. and Vandebriel, R. J. (2016). State of the art in non-animal approaches for skin sensitization testing: From individual test methods towards testing strategies. *Arch Toxicol* 90, 2861-2883. doi:10.1007/s00204-016-1842-4
- Fentener van Vlissingen, M., Grootaers, G., Groothuis, G. et al. (2017). Vital tissue: A fresh human tissue supply chain to enable translation research recommendations for the sustainable use of vital human tissue material for (applied) scientific research to reduce the need for animal testing. *TNO*. <http://resolver.tudelft.nl/uuid:dfbd9986-6fce-4efa-9a2f-859e24f4ae15>
- Gouliarmou, V., Lostia, A. M., Coecke, S. et al. (2018). Establishing a systematic framework to characterise in vitro methods for human hepatic metabolic clearance. *Toxicol In Vitro* 53, 233-244. doi:10.1016/j.tiv.2018.08.004
- Grosberg, A., Alford, P. W., McCain, M. L. et al. (2011). Ensembles of engineered cardiac tissues for physiological and pharmacological study: Heart on a chip. *Lab chip* 11, 4165-4173. doi:10.1039/c1lc20557a
- Hassan, I., El-Masri, H., Kosian, P. A. et al. (2017). Neurodevelopment and thyroid hormone synthesis inhibition in the rat: Quantitative understanding within the adverse outcome pathway framework. *Toxicol Sci* 160, 57-73. doi:10.1093/toxsci/kfx163
- Huh, D., Matthews, B. D., Mammoto, A. et al. (2010). Reconstituting organ-level lung functions on a chip. *Science* 328, 1662-1668. doi:10.1126/science.1188302
- ICCVAM – Interagency Coordinating Committee on the Validation of Alternative Methods (2018). A Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medicinal Products in the United States.
- Kim, H. J. and Ingber, D. E. (2013). Gut-on-a-chip microenvironment induces human intestinal cells to undergo villus differentiation. *Integr Biol (Camb)* 5, 1130-1140. doi:10.1039/c3ib40126j
- Kim, S. H., Chi, M., Yi, B. et al. (2014). Three-dimensional intestinal villi epithelium enhances protection of human intestinal cells from bacterial infection by inducing mucin expression. *Integr Biol (Camb)* 6, 1122-1131. doi:10.1039/c4ib00157e
- Lauschke, V. M., Vorrink, S. U., Moro, S. M. L. et al. (2016). Massive rearrangements of cellular microRNA signatures are key drivers of hepatocyte dedifferentiation. *Hepatology* 64, 1743-1756. doi:10.1002/hep.28780
- Lee, K. K., McCauley, H. A., Broda, T. R. et al. (2018). Human stomach-on-a-chip with luminal flow and peristaltic-like motility. *Lab Chip* 18, 3079-3085. doi:10.1039/C8LC00910D
- Marx, U., Andersson, T. B., Bahinski, A. et al. (2016). Biology-inspired microphysiological system approaches to solve the prediction dilemma of substance testing. *ALTEX* 33, 272-321. doi:10.14573/altex.1603161
- Maschmeyer, I., Lorenz, A. K., Schimek, K. et al. (2015). A four-organ-chip for interconnected long-term co-culture of human intestine, liver, skin and kidney equivalents. *Lab Chip* 15, 2688-2699. doi:10.1039/C5LC00392J
- Mastrangeli, M., Millet, S., The ORCID partners et al. (2019). Organ-on-chip in development: Towards a roadmap for organs-on-chip. *ALTEX* 36, 650-668. doi:10.14573/altex.1908271
- McMullen, P. D., Andersen, M. E., Cholewa, B. et al. (2018). Evaluating opportunities for advancing the use of alternative methods in risk assessment through the development of fit-for-purpose in vitro assays. *Toxicol In Vitro* 48, 310-317. doi:10.1016/j.tiv.2018.01.027
- Norona, L. M., Nguyen, D. G., Gerber, D. A. et al. (2016). Editor's highlight: Modeling compound-induced fibrogenesis in vitro using three-dimensional bioprinted human liver tissues. *Toxicol Sci* 154, 354-367. doi:10.1093/toxsci/kfw169
- O'Neill, A. T., Monteiro-Riviere, N. A. and Walker, G. M. (2008). Characterization of microfluidic human epidermal keratinocyte culture. *Cytotechnology* 56, 197-207. doi:10.1007/s10616-008-9149-9
- OECD (2014). Guidance Document for Describing Non-Guideline In Vitro Test Methods. *Series on Testing and Assessment No. 211*. OECD Publishing Paris. doi:10.1787/9789264274730-en
- OECD (2016a). Guidance Document for the Use of Adverse Outcome Pathways in Developing Integrated Approaches to Testing and Assessment (IATA). *Series on Testing and Assessment No. 260*. OECD Environment, Health and Safety Publications. [https://one.oecd.org/document/ENV/JM/MONO\(2016\)67/en/pdf](https://one.oecd.org/document/ENV/JM/MONO(2016)67/en/pdf)
- OECD (2016b). Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for Skin Sensitisation. *OECD Series on Testing and Assessment No. 256*. OECD Publishing, Paris. doi:10.1787/9789264279285-en
- 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
- Oleaga, C., Bernabini, C., Smith, A. S. et al. (2016). Multi-organ toxicity demonstration in a functional human in vitro system composed of four organs. *Sci Rep* 6, 20030. doi:10.1038/srep20030
- Piersma, A. H., Burgdorf, T., Louekari, K. et al. (2018a). Workshop on acceleration of the validation and regulatory acceptance of alternative methods and implementation of testing strategies. *Toxicol In Vitro* 50, 62-74. doi:10.1016/j.tiv.2018.02.018
- Piersma, A. H., van Benthem, J., Ezendam, J. et al. (2018b). Validation redefined. *Toxicol In Vitro* 46, 163-165. doi:10.1016/j.tiv.2017.10.013
- Pittman, M. E., Edwards, S. W., Ives, C. et al. (2018). AOP-DB: A database resource for the exploration of adverse outcome pathways through integrated association networks. *Toxicol Appl Pharmacol* 343, 71-83. doi:10.1016/j.taap.2018.02.006
- Pridgeon, C. S., Schlott, C., Wong, M. W. et al. (2018). Innovative organotypic in vitro models for safety assessment: Aligning with regulatory requirements and understanding models of



- the heart, skin, and liver as paradigms. *Arch Toxicol* 92, 557-569. doi:10.1007/s00204-018-2152-9
- RIVM (2018). Roadmap for animal-free innovations in regulatory safety assessment. National Institute for Public Health and the Environment (RIVM).
- Scialli, A. R., Daston, G., Chen, C. et al. (2018). Rethinking developmental toxicity testing: Evolution or revolution? *Birth Defects Res* 110, 840-850. doi:10.1002/bdr2.1212
- Shim, K. Y., Lee, D., Han, J. et al. (2017). Microfluidic gut-on-a-chip with three-dimensional villi structure. *Biomed Microdevices* 19, 37. doi:10.1007/s10544-017-0179-y
- Vedula, E. M., Alonso, J. L., Arnaout, M. A. et al. (2017). A microfluidic renal proximal tubule with active reabsorptive function. *PLoS One* 12, e0184330. doi:10.1371/journal.pone.0184330
- Wagner, I., Materne, E.-M., Brincker, S. et al. (2013). A dynamic multi-organ-chip for long-term cultivation and substance testing proven by 3D human liver and skin tissue co-culture. *Lab Chip* 13, 3538-3547. doi:10.1039/c3lc50234a
- Xiao, S., Coppeta, J. R., Rogers, H. B. et al. (2017). A microfluidic culture model of the human reproductive tract and 28-day menstrual cycle. *Nat Commun* 8, 14584. doi:10.1038/ncomms14584
- Yu, J., Peng, S., Luo, D. et al. (2012). In vitro 3D human small intestinal villous model for drug permeability determination. *Biotechnol Bioeng* 109, 2173-2178. doi:10.1002/bit.24518
- Zhang, B. and Radisic, M. (2017). Organ-on-a-chip devices advance to market. *Lab Chip* 17, 2395-2420. doi:10.1039/C6LC01554A

### Conflict of interest

The authors declare that they have no conflicts of interest.

### Acknowledgements

We would like to thank the workshop participants (in alphabetic order): Martin Bennink, Flemming Cassee, Susana Chuva de Sousa Lopes, Samuel Constant, Wim de Jong, Sylvia Le Dévédec, Christopher Drewell, Sue Gibbs, Betty Hakkert, Frieke Kuper, Henriette Lanz, Roos Masereeuw, Andre Muller, Eduard Pelssers, Aldert Piersma, Mieke Schutte, Adrienne Sips, Peter Theunissen, Stefan Vaessen, Janny van den Eijnden, Jan Willem van der Laan, Andries van der Meer, Meike van der Zande, Evita van de Steeg, Jacqueline van Engelen, and Remco Westerink. Furthermore, Flemming Cassee, Wim de Jong and Adrienne Sips are acknowledged for providing significant input into the workshop. This work was supported by the Dutch Ministry of Agriculture, Nature and Fishery.