Archives of Toxicology
https://doi.org/10.1007/s00204-024-03922-z

REVIEW ARTICLE

Bridging imaging-based in vitro methods from biomedical research to regulatory toxicology

Monica Piergiovanni[1] · Milena Mennecozzi[1] · Erio Barale-Thomas[2] · Davide Danovi[3] · Sebastian Dunst[4] · David Egan[5] · Aurora Fassi[1] · Matthew Hartley[6] · Philipp Kainz[22] · Katharina Koch[7,8] · Sylvia E. Le Dévédec[9] · Iris Mangas[10] · Elena Miranda[11] · Jo Nyffeler[12] · Enrico Pesenti[13] · Fernanda Ricci[14] · Christopher Schmied[15] · Alexander Schreiner[16] · Nadine Stokar-Regenscheit[17] · Jason R. Swedlow[18] · Virginie Uhlmann[19] · Fredrik C. Wieland[20] · Amy Wilson[21] · Maurice Whelan[1]

Received: 15 October 2024 / Accepted: 26 November 2024

Abstract
Imaging technologies are being increasingly used in biomedical research and experimental toxicology to gather morphological and functional information from cellular models. There is a concrete opportunity of incorporating imaging-based in vitro methods in international guidelines to respond to regulatory requirements with human relevant data. To translate these methods from R&D to international regulatory acceptance, the community needs to implement test methods under quality management systems, assess inter-laboratory transferability, and demonstrate data reliability and robustness. This article summarises current challenges associated with image acquisition, image analysis, including artificial intelligence, and data management of imaging-based methods, with examples from the developmental neurotoxicity in vitro battery and phenotypic profiling assays. The article includes considerations on specific needs and potential solutions to design and implement future validation and transferability studies.

Keywords Imaging-based in vitro methods · Regulatory toxicology · High content imaging · Artificial intelligence · Validation · 3Rs

Introduction

Imaging technologies are playing a central role in biomedical sciences, serving as a powerful tool to investigate a variety of responses in cell-based assays. The term imaging-based in vitro methods here refers to test methods or assays that use imaging technologies, for example light/fluorescence microscopy and high-content imaging (HCI), to quantify morphological and functional properties in cellular models. The primary output of imaging-based methods are digital images that can have different dimensions and content (*e.g.*, 2D images, 3D volumes, time-lapse). However, these images are rarely the final readout of the assay, as it is typically necessary to process images and extract image features for further (quantitative) analysis. These analysis, performed using open-source or proprietary software that can include machine/deep learning (ML/DL) algorithms, lead to the evaluation of a wide range of tissue and cellular properties, including cell survival, size, shape, morphology, protein expression levels, and subcellular localisation of organelles and compounds. These parameters are used to gather information on a biological effect, usually in response to external stimuli (*e.g.*, compound exposure, genetic modification). Imaging can also provide information on the physiological state or function of a biological system. Functional read-outs or endpoints cover features such as cell migration, calcium flux, contractility, and angiogenesis. Analysing these features can help in getting valuable insights into cellular processes, identifying potential drug targets, and assessing the safety and efficacy of exogenous compounds.

Recently, there is increasing interest in regulatory toxicology to accelerate the transition of human-relevant, imaging-based in vitro methods into testing strategies for decision-making. Regulatory toxicology allows authorities to take safety decisions for the protection of human health and the

Monica Piergiovanni and Milena Mennecozzi have contributed equally.

Extended author information available on the last page of the article

Published online: 13 February 2025

environment against harmful effects of various substances (Schwenk et al. 2002). However, international guidance documents and guidelines do not yet incorporate imaging-based in vitro methods, hindering a widespread application for regulatory purposes and thus the transition to non-animal methods, following the principle of the 3Rs (Refinement, Reduction, Replacement) (Russell, Burch, and Hume 1959). Novel technologies around in vitro models are evolving, but there is scarce information on simple and efficient methods to evaluate microtissue (*i.e.,* organoid) morphology. Application of advanced histotechnologies and digital pathology workflow as used in regulatory toxicological pathology will allow future correlation of preclinical in vivo data and translational biomarkers to the one generated in vitro. Ideally, interdisciplinary teams, as biologists, bioengineers, toxicologists, and pathologists, work together on implementation of novel morphology-based readouts (Stokar-Regenscheit et al. 2024).

This article examines the main challenges to the adoption of imaging-based in vitro methods for regulatory purposes. It outlines the state of imaging technologies across different regulated fields, and analyses the key steps of the imaging process, providing platform-agnostic recommendations to both researchers and regulators. Special focus is given to evaluating transferability and reproducibility as a critical step in method qualification and/or validation and to ensuring data quality and transparency, with the aim to increase confidence in the data generated and enabling its practical use in decision-making.

Use of imaging technologies in medical regulated fields

Defining standardised processes and procedures is a common approach across many regulated sectors, agencies, and jurisdictions and it is already applied to imaging technologies. In the medical field, imaging is frequently used in clinical trials to support approval of drugs and biological products. The U.S. Food and Drug Administration (FDA) issued the "Clinical Trial Imaging Endpoint Process Standards Guidance for Industry" with the goal of assisting sponsors in optimising the quality of data generated with imaging, by discussing aspects related to the whole imaging process, including acquisition, visualisation, storage, archiving, and interpretation of image data (U.S. Food and Drug Administration 2018).

Describing processes to guarantee the quality of an instrumental measurement, thus including microscopes and other imaging platforms, is an internationally agreed approach, as confirmed by the extensive use of documents such as the "Validation of analytical procedures" and the "Analytical Procedure Development" issued by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) and adopted by many countries in the world. The Q2(R2)/Q14 guideline, for instance, describes the principles relating to the development and maintenance of the processes used for analytical procedures and the respective validation tests to be implemented (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 2023a; 2023b).

Histopathology data are essential for toxicity studies, evaluating drug-induced toxic effects in laboratory animals and its relevance for human safety, as required by regulatory authorities (Mehrvar et al. 2021). Digital pathology employs digital images and data analysis to diagnose and monitor diseases from tissue samples. Digital toxicological histopathology has been broadly adopted in preclinical compound development for informal consultation and peer review (Boisclair et al. 2022). Robust regulatory frameworks governing the use of data generated by digital pathology are already in place and the community has extensive experience in standardising staining procedures and utilising imaging data in regulatory dossiers. Recent initiatives provide valuable examples. By ensuring strict GDPR and access control, Bigpicture[1] offers a unified infrastructure with the tools needed to exploit the data gathered (3 million curated, quality-controlled human and animal digital slides), that can be used for research and as reference/historical control (Moulin et al. 2021; Rudmann et al. 2023; Jacobsen et al. 2021). The regulatory landscape for digital pathology is evolving rapidly, as more countries and regions recognise the potential benefits of this technology for public health and innovation. In diagnostic pathology, several digital pathology devices and software for primary diagnosis of some cancers, such as breast, prostate, and gastrointestinal, have been approved by the FDA in the United States since 1995.[2] In 2016, the FDA also provided recommendations on the technical performance assessment that should be submitted for regulatory evaluation of a digital system, through publication of guidance documents (FDA 2023; 2016). Starting from May 2022, the European Commission has enforced a new regulation on in vitro diagnostic medical devices (IVDR). Digital pathology devices and software are considered high-risk products under the IVDR, and as such need to meet higher standards of clinical evidence and post-market surveillance. In other regions, such as Asia–Pacific, Latin America, and Africa, the regulatory status of digital pathology varies widely, depending on the level of development, infrastructure, and resources of each country. Some

[1] https://bigpicture.eu/

[2] https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

countries, such as Japan, China, and India, have established or are developing their own regulatory frameworks for digital pathology, while others, such as Brazil, Mexico, and South Africa, are following or adapting the international standards and guidelines (Barisoni et al. 2020). These recent advances in Whole Slide Imaging (WSI) have created an enormous opportunity to provide more quantitative, objective, and consistent assessments of pathology datasets, improve workflow efficiency, prevent slide deterioration, and develop decision support systems. However, there are also challenges and barriers to overcome, such as the lack of harmonised standards, guidelines, and validated methods, as well as the need for robust data security and quality assurance (Zuraw and Aeffner 2022). Virtual staining is now emerging as an innovative technology which is facilitating data re-use, improving safety decision-making, streamlining workflows, and ultimately promoting standardisation. Virtual staining overcomes the limitations of manual staining workflows (time consuming, staining variability, small tissue size) by using computational algorithms to generate digital representations of various staining techniques on unstained sections (Rivenson et al. 2019).

Current status of imaging technologies for safety testing of chemicals in OECD TGs

An OECD Test Guideline (TG) is an accepted international standard test method for safety testing of chemicals in different sectors (*e.g.,* industrial chemicals, medicinal products, cosmetics, food additives, pesticides). A TG can be used to obtain a range of information on a test item, from physiochemical properties to biological adverse effects. It includes a relevant test system (biological model) for a toxicological endpoint to be measured and a data interpretation procedure or prediction model to derive results and draw conclusions. Often, the term assay is also used interchangeably to refer to a test method. A test battery is a combination of methods designed to complement each other to investigate a toxicological endpoint (OECD 2005).

A key step for regulatory implementation of a new test method is the possibility to implement it under a quality assurance scheme, with Good Laboratory Practices (GLP) being the most common one. Experimental data generated using OECD TGs that are performed under GLP are covered by the OECD Mutual Acceptance of Data (MAD) system. This means these data are accepted in all OECD and MAD-adherent countries and used in many different hazard and risk assessment contexts related to human health and environment protection. This has significant benefits since it avoids unnecessary duplication of testing, resulting in appreciable financial savings and reduction in animal use. GLP is applicable to non-clinical health and environmental studies and ensures that the quality and integrity of data can be demonstrated, enabling study reconstruction. In OECD countries, the application of GLP is legally binding for the safety assessment of chemicals and chemical products proposed for authorisation. Monitoring authorities confirm GLP compliance of test facilities and studies through inspections. GLP advisory and consensus documents, very relevant for HCI, are the application of GLP principles to in vitro studies (No. 14), the application of GLP principles to computerised systems (No. 17), and GLP data integrity (No. 22). By adhering to these principles, HCI studies can generate data suitable for health and environmental safety assessment.

To be adopted as an OECD TG, a test method needs to be validated according to OECD Guidance Document 34 (GD34) (OECD 2005), demonstrating the relevance and reliability of the method, as well as establishing its performance against reference materials. GD34 describes the various approaches to validation and provides indication on how to perform a validation study. Additionally, it addresses the need for peer review of the results and the subsequent international acceptance of data derived from validated tests for regulatory purposes. Currently under revision, GD34 aims to incorporate updated guidance for emerging technologies, among others.

Two promising methods using HCI are getting closer to regulatory use and international acceptance: *i.* the developmental neurotoxicity (DNT) in vitro battery, and *ii.* Phenotypic Profiling (*i.e.*, Cell Painting assay).

Developmental neurotoxicity in vitro battery DNT refers to a change in the structure or function of the nervous system after exposure to a chemical during the pre- and/or postnatal period. The current testing paradigm for DNT mainly relies on results on results from in vivo guideline studies (OECD TG 426 (OECD 2007), OECD TG 443 with DNT cohort (OECD 2018a)). However, the limited in vivo DNT testing, the high costs, the species-specific peculiarities of brain development, and the commitment to reduce animal use, have spurred the development of alternative, non-animal methods to screen chemicals for DNT potential.

Recently, a DNT in vitro battery (DNT IVB) has been assembled (Sachana, Shafer, and Terron 2021; Harrill et al. 2018; Blum et al. 2023), comprising of multiple test methods assessing individual key neurodevelopmental processes (KNDPs). Some methods rely on HCI and are combined with Artificial Intelligence (AI)-based cell annotation. For instance, the "Human Neurosphere Assay" (Koch et al. 2022) assesses chemical effects on the proliferation and differentiation of human foetal neural progenitor cells cultured as neurospheres in vitro. Plating differentiating neurospheres on an artificial extracellular matrix causes radial migration of differentiating cells out of the sphere core and the formation of a circular migration area, in which the different cell types can be quantified using

image analysis. The migration distance of radial glia cells, neurons and oligodendrocytes, the differentiation and morphology of neurons, and the differentiation of oligodendrocytes is assessed based on brightfield images and immunocytochemical staining using the image analysis software 'Omnisphero' (Schmuck et al. 2017; Förster et al. 2022) and two Convolutional Neural Networks (CNN).

The establishment of the DNT IVB was supported by the European Food Safety Authority (EFSA) which funded the assessment of the DNT potential of 120 substances, including several chemicals of DNT concern, such as pesticides. Within the scope of this screening approach, a performance analysis using human DNT positives and negatives as well as positives from DNT in vivo guideline studies confirmed high reproducibility and accuracy of the DNT IVB results, while also identifying areas for improvement. The integration of test methods identifying glia cell-specific toxicities (*i.e.,* astrocytes, radial glia, and microglia) and the use of exclusively human-based cell systems shall improve the DNT IVB performance in the future (Masjosthusmann et al. 2020; Serafini et al. 2024). The cell systems within the DNT IVB have undergone rigorous characterisation and mechanistic assessment. Multiple key steps are fundamental for transitioning in vitro DNT test methods to regulatory use:

- Endpoint selection: Endpoints must be scientifically validated to ensure the cell systems accurately represent human biology, respond to physiological stimuli, and are sensitive to known human DNT-positive chemicals (Koch et al. 2022).
- Test method design: Predictive test methods need to include endpoint-specific controls on each test plate, ensure an appropriate dynamic range, and enable the calculation of concentration–response relationships. Moreover, test methods should have sufficient throughput and be applicable for screening purposes (Crofton et al. 2011).
- Determination of test method readiness: Within the process of test method optimisation and validation, the readiness of the test method needs to be evaluated (Bal-Price et al. 2018).
- Comprehensive chemical testing: OECD proposed a list of approximately 100 chemicals for use as DNT reference chemicals (OECD 2023a). This common set of chemicals served as a foundation for comparing current DNT IVB performance against the available in vivo DNT studies and evaluating the added value of possible new assays developed for inclusion in the DNT IVB.
- Establishment of positive and negative control compounds with known mechanism of actions in the specified biological system and modelled KNDP (*i.e.,* endpoint specific controls, training set (Martin et al. 2022; Mundy et al. 2015)).

Data generated using the DNT IVB has been used for hazard identification in EFSA-coordinated case studies on pesticides such as deltamethrin, acetamiprid, and flufenacet. (OECD 2023a; EFSA et al. 2024). To fulfil the regulatory requirements for the implementation of the DNT IVB in OECD test guidelines, efforts to ensure the transferability of the DNT IVB test methods are ongoing.

Phenotypic profiling (PP) Phenotypic profiling is an emerging in vitro method in which cells are fluorescently labelled to visualise the morphological changes that a perturbagen (such as a chemical compound or a genetic modification) induces. This method assumes that morphological changes occurring at the cellular level are an indicator for functional perturbations of tissues and organs. The term "profiling" is in contrast to targeted phenotypic assays such as the E-Morph Screening Assay (Klutzny et al. 2022), which give information on a pre-defined process. Phenotypic profiling aims to describe cell morphology in an encompassing way. For this purpose, the generated microscopy images are used to extract many different image features using Machine Learning (ML) approaches based on handcrafted features and more recently using DL (Caicedo et al. 2017; Chandrasekaran et al. 2021). Comparison of the measured profiles to phenotypic profiles of known reference compounds allows the identification of mechanisms of actions (MoA). Therefore, it also enables the investigation of potential adverse effects, without requiring *a priori* knowledge of the involved cellular targets. The large volumes of image data generated using PP provide great opportunities to predict compound activities for human-relevant disease endpoints and to identify the underlying MoA using ML and, importantly, independent from animal experiments (Odje et al. 2024). However, these data also pose great challenges in ensuring tractability and transparency of image and data analysis routines as well interpretation of the phenotypic profiles in a regulatory context.

The concept of PP was pioneered by the laboratory of A.E. Carpenter at the Broad Institute of Harvard and MIT (Gustafsdottir et al. 2013; Bray et al. 2016) and has quickly been successfully implemented, optimised, and standardised for high-throughput phenotypic profiling at various sites worldwide (Cimini et al. 2023; Tromans-Coia et al. 2023). This broadly used PP method is often referred to as "Cell Painting" (CP). Comparative studies using control compound plates at multiple sites have shown that CP is a robust and reproducible method with high and cross-platform interoperability, even if each laboratory needed to adjust some assay parameters (Cimini et al. 2023; Tromans-Coia et al. 2023). Sources of variability originate from cell culture

conditions (*e.g.,* cell density at the imaging stage), solvent percentage (*e.g.,* different amount of DMSO), staining procedure, (*e.g.,* dye signal stability over time), and finally from the use of different microscopy systems and data analysis software and approaches (Arevalo et al. 2024). CP has been used in large consortia like JUMP (Joint Undertaking for Morphological Profiling)[3] or OASIS (Omics for Assessing Signatures for Integrated Safety),[4] in which academic institutes, government agencies, industry organisations, and non-profit organisations are collaborating to create the worldwide largest CP-based image dataset and phenotypic profiles for hundreds of thousands of small molecules, industrial/agricultural chemicals or genetic perturbations, which will eventually become publicly available as resource datasets (Chandrasekaran et al. 2024; Wolff et al. 2024). Furthermore, CP already demonstrated its applicability in the context of hazard assessment of industrial chemicals. In vitro points-of-departures derived from human cells with CP were similar or more conservative than in vivo animal points-of-departure for 68% of the chemical tested (Nyffeler et al. 2020). Over 1000 chemicals have been tested with this method so far, and the phenotypic profiles are available in the public EPA CompTox Chemicals Dashboard[5] (Nyffeler et al. 2023). As a next step, standard CP would need to formally demonstrate their applicability for regulatory purposes in comprehensive validation studies, a first key step towards their implementation as an OECD TG.

Imaging-based in vitro methods towards regulatory use

A successful imaging-based in vitro method needs to be based upon a reliable and relevant biological model. Considerations on the choice of the best biological model for various purposes of regulatory interest are out of the scope of this work. Many resources discuss technical and biological variability, concordance with in vivo data, and frameworks for building confidence in in vitro methods for risk assessment and regulatory purposes (National Academies of Sciences and Medicine 2023; Senft et al. 2023; Negro et al. 2018; Pistollato et al. 2021). Focusing exclusively on technology, this article analyses key steps of the imaging process, as described in Table 1. Downstream data analysis of parameters extracted from images is not in the scope of this work.

[3] https://jump-cellpainting.broadinstitute.org/

[4] https://oasisconsortium.org/

[5] https://comptox.epa.gov/dashboard/

Transferability and reproducibility

Obtaining regulatory acceptance for imaging-based in vitro methods necessitates careful assessment of transferability and reproducibility of the results within and between laboratories, as required for the test method validation in GD34 (OECD 2005). Transferability studies evaluate the performance of a test method in a naïve laboratory by measuring specificity and sensitivity of the test method against a set of reference compounds and comparing these data with historical values generated in the developer's laboratory. However, significant challenges exist, mainly due to discrepancies in instruments, protocols, reagents, biological systems, and human error. Incentive grants and publications in academia, while valuable, often do not directly incentivise transferability and validation studies. Advocating for more pre-competitive spaces is therefore crucial to overcome this hurdle. Harmonising imaging terminology is a first effort that would facilitate setting up of transferability studies and allow transparent performance evaluation of test methods implemented in different laboratories.

Developing robust imaging-based in vitro methods requires careful consideration of several factors. Researchers should prioritise well-characterised and readily available biological models. Well-defined, quantifiable and measurable endpoints, like the number of cells positive for a specific type or functional marker, are also preferable. Multiplexing imaging readouts within the assay design provides a more comprehensive understanding of cellular phenotypes and their interaction, especially in response to stimuli, while increasing data throughput. Finally, it is crucial to strike a balance between assay complexity and both transferability and biological relevance. Overly complex assays may be difficult to transfer across laboratories, while too simplistic assays may not fully capture the biological process under investigation. By carefully weighing these factors, researchers can design robust imaging-based in vitro methods that generate meaningful and reproducible results, ultimately advancing the use of imaging. Using public repositories of cells and reagents guarantees a consistent source of materials with homogeneous properties, mitigating batch-to-batch variability in transferability studies. Utilising curated lists of reference chemicals and implementing rigorous quality control steps throughout the workflow, encompassing both cellular model setup and the imaging process, are crucial for maintaining consistency and evaluating inter-experiment variability. A balance between sensitivity and specificity (around 80% for both) is usually accepted for regulatory decision-making, which enables to have a reasonable number of false positive and false negatives (Kolle et al. 2015).

Planning for transferability starts during assay development, when scientists are recommended to prioritise automated procedures and implement a clear documentation

Table 1 Step-by-step recommendations along the imaging workflow to facilitate the transferability and quality management of imaging-based in vitro methods

Image acquisition	
Instrument calibration	Mechanical calibration of the microscope stage is particularly important for experiments that involve time-lapse imaging, where the measurement needs to be taken in the exact same location at multiple time points. When fluorescence is used, calibration of the fluorescent intensity with standardised materials (polymer or glass beads) is recommended, especially for quantitative measurements (Halter et al. 2014), (Kwee et al. 2021). A list of suggested calibration metrics can be found in (Huisman et al. 2021).
Instrument configuration	The configuration of the instrument should be accurately reported for the specific test method. For microscopy applications, common parameters include illumination, magnification, excitation and emission wavelengths, objective lenses, and type of detector. Specific technical parameters for the different imaging technologies can be found in (Sarkans et al. 2021; Nelson et al. 2021; Montero Llopis et al. 2021).
Instrument performance	The instrument performance should be assessed to establish sensitivity, specificity, and dynamic ranges of the instrument, over time and at varying environmental (*i.e.,* humidity and temperature) conditions. Specific parameters for the different imaging technologies can be found in (Sarkans et al. 2021; Nelson et al. 2021; Montero Llopis et al. 2021).
Image analysis	
Data formats	The use of standardised open data formats like OME.tiff (Hammer et al. 2021), OME-Zarr (Moore et al. 2021), DICOM[a] (mainly used in digital pathology) should be adopted across organisations and geographical boundaries. These image formats retain the metadata and ensure comparability of data across different platforms and studies, enabling re-analysis of image data (Allan et al. 2012; Hartley et al. 2022). Long-term compatibility of these formats to future hardware and software should also be considered.
Algorithm performance	The performance of the ML/DL model used for image analysis should be accurately measured with problem-specific metrics, some guidance on different approaches and recommendations can be found in (Laine et al. 2021; Maier-Hein et al. 2024; McGenity et al. 2024). Open and transparent ML model training documentation, along with version control of both training data and models, can accelerate the development, validation, and implementation of robust ML tools for image analysis. Exchanges with established ISO technical committees (*e.g.,* ISO/IEC JTC 1/SC 42 – Artificial Intelligence) would be beneficial to reach international agreement on specific performance standards.
Software verification	Image analysis verification can be done with a fixed set of input images and by proving that the results obtained are consistent. When using DL algorithms, caution must be exercised to appropriately define the range of applicability of a model and avoid domain shift (*i.e.,* discrepancies in statistics between training and test data).
Metadata	
Metadata characteristics	Metadata should include detailed descriptions of the experimental procedures, characteristic of biological samples (*i.e.,* source, preparation, treatments), microscope hardware specifications (*i.e.,* model, camera), image acquisition settings (*i.e.,* magnification, resolution, illumination, exposure time, filter used), image analysis methods and (hyper-)parameters, data annotation procedures, detailing instrument performance and calibration metrics (*i.e.,* quality control), and analysed data. Existing guidelines provide lists of relevant parameters, for instance: REMBI: Recommended Metadata for Biological Images, MIHCSME for high-content imaging data (Hosseini et al. 2023), 4DN-BINA-OME Microscopy Metadata Specifications (Hammer et al. 2021; Sarkans et al. 2021; Huisman et al. 2021).
Metadata annotation	Scrutinising data collection and annotation is vital for model performance and dataset reliability. It is suggested to focus on label accuracy, as incorrect labels can lead to critical errors, and aim for class balance to prevent model bias toward overrepresented categories. Annotator bias can be avoided by validating work from multiple annotators. Capturing comprehensive metadata throughout the experimental workflow remains a challenge, as highlighted in a recent survey (Schmidt et al. 2022). Potential solutions include utilising electronic laboratory notebooks, while another option is to leverage instrument-integrated metadata annotation tools to minimise human error and enhance metadata quality and consistency (Rigano et al. 2021; Kunis et al. 2021; Ryan et al. 2021). For GLP compliance, metadata should be preferably reported automatically by the instrument or manually entered during the experiment to ensure data integrity. Retrospective metadata addition is not acceptable.
Ontologies	An ontology acts as a shared vocabulary, providing a standardised framework for describing concepts, their relationships and the properties of those concepts. Thus, a well-designed ontology serves as a common language for describing and annotating image data, ensuring consistent data capture and reporting, regardless of the specific instrument or user. Existing initiatives, such as the Ontology Lookup Service (OLS)[b], represents a promising solution for identifying and implementing biomedical ontologies. Harmonised and user-friendly ontologywill facilitate the re-use of imaging data across platforms and operators. The use of controlled terminologies (e.g. INHAND[c], SNOMED[d]) is advised to limit the risk of spelling errors and multiplication of synonyms which hinders searches.

Table 1 (continued)

Data management	
Data repositories	Data repositories should be designed to accommodate both processed and raw imaging data, to allow for an independent and comprehensive analysis and evaluation of findings by peer scientists and regulators, as well as to guarantee traceability for GLP audit purposes. Concerns exist regarding adhering to GLP requirements with cloud-based storage, thus current practices often involve a hybrid approach, maintaining a physical archive alongside cloud storage for everyday use. In GLP archives, the raw data should be archived together with the metadata and not separated. Examples of existing image repositories can be found here (OMERO (Allan et al. 2012; Burel et al. 2015; Li et al. 2016), the BioImage Archive[e] (Hartley et al. 2022), IDR (Williams et al. 2017), and PIDAR[f]). Some technologies or configurations, such as HCI or 3D/time-lapse acquisitions, routinely generate large datasets for each experiment, requiring high computational power and scalable storage.
Data management plan	The data management plan ensures transparency, reproducibility, and the long-term value of scientific research. It should: • include a general description of the project/experiment and clearly defined roles and responsibilities for data collection, curation, storage, and access; • include a list of all the data types expected to be generated and a comprehensive strategy for documenting the data formats, file naming conventions, and associated metadata; • outline a plan for secure and reliable data storage and transmission throughout the project duration and beyond; • specify data access policies, including who will have access to data, right to modify it or visualise it; • address any ethical or legal concerns associated with data collection, storage, and sharing.
Data integrity and security	Data integrity measures should be in place to protect data from manipulation, corruption or loss over time, including long-term archiving of raw data and their metadata. While cloud storage offers advantages for remote access and collaboration in imaging, robust security measures are essential to ensure patient or client privacy. Copies of raw data and raw images shared with the scientific community (*e.g.* in data repositories) can be separated from mapping files and metadata containing confidential data to enhance the level of security.

[a]https://www.dicomstandard.org/about-home

[b]https://www.ebi.ac.uk/ols4

[c]https://www.inhand.com/en/

[d]https://www.snomed.org/

[e]https://www.ebi.ac.uk/bioimage-archive/

[f]https://idr.openmicroscopy.org/

management strategy, including detailed study plans and Standard Operating Procedures (SOPs). Study plans for imaging-based in vitro methods should be platform-agnostic, sufficiently detailed and complete to facilitate efficient and effective reproducibility and transferability of experimental results. While setting up the transferability study, it is crucial for the scientists to analyse the SOPs and adapt them to the receiving laboratory. Detailed SOPs should encompass all implementation details of the test method preparation and set up, image acquisition and analysis pipelines, and data management. Before starting an imaging experiment, proper instrument calibration should be performed, ensuring that instruments are performing according to specifications. Unfortunately, systematic instrument calibration is often overlooked, with many researchers relying on the acquired images themselves to set the acquisition parameters for each experiment. This practice significantly hinders efficient transferability across sites and even between different operators within the same site. A more reliable approach involves calibration with standardised materials available for the fluorescent ranges commonly used in biomedical research (Halter et al. 2014; Kwee et al. 2021).

Measurable instrument performance enhances the reliability of the test methods, thus maximising the concordance of the results across sites. Key parameters to consider include uniform illumination of the field of view, precise excitation and emission intensities, and optimal focus settings for high-quality images. However, in some microscopy areas such as HCI, manufacturers are reluctant to disclose details on the instrument settings and controls and prefer to adapt their products to the specific needs of the customer ona case-by-case basis. This lack of transparency implies that only few imaging-based in vitro methods would be used for regulatory toxicology. Currently, there is a growing tendency for test methods that can be implemented on a single instrument within a single facility to be accepted as OECD TGs, effectively establishing them as *de facto* standards. To mitigate the risk of test method monopolisation and streamline the identification of suitable testing facilities, openly sharing instrument performances based on agreed technical parameters, along with reagents details, is crucial. Moreover, precise specifications of the image acquisition parameters, including magnification, illumination conditions, fluorescence channel selection, image resolution, and any specific settings required

for the chosen imaging modality, must be documented in the SOPs. A well-defined image analysis workflow should be outlined, ideally accompanied with code (Miura and Nørrelykke 2021), and should include details on the image analysis algorithms and parameters used (*e.g.,* background correction, noise reduction, segmentation, and quantification). Finally, it should be specified how the data are reported: file format, data normalisation procedures, raw and processed images.

Recognising a critical need for standardisation in imaging technologies, organisations such as Global BioImaging[6] and QUAREP-LiMi[7] (Quality assurance and reproducibility for instrument and images in light microscopy) facilitate best practices and collaboration across stakeholders (Nelson et al. 2021; Schmied et al. 2024). A successful example of collaboration between QUAREP-LiMi and major camera manufacturers led, for instance, to the definition of a common vocabulary for describing scientific camera performance (Marx 2022). Another crucial area is standardisation of microplates used for cellular imaging, as current standards inadequately address plate bottom flatness. This inevitably causes problems in systematically adjusting the focus, an important issue for HCI, where thousands of wells are automatically imaged. To address this issue, ISO (International Organization for Standardization) and ANSI (American National Standards Institute) should coordinate efforts to revise and update plate standards. Additionally, international initiatives like Euro-BioImaging[8] are focused on optimising in vitro imaging data collection providing access to high-end instrumentation and fostering collaboration across borders.

Importantly, for transferability studies of methods that use advanced imaging technologies, the receiving laboratory should be selected based on presence of suitable equipment and specialised personnel with advanced training in microscopy, imaging, and data analysis. Establishing multiple centralised imaging testing facilities across the globe is a proposed solution to speed up validation studies and ensure the generation of reproducible data. With a similar reasoning, the Tox21 consortium was founded by a group of US agencies with the goal of using high throughput screening to generate data on data-poor compounds for safety assessment (Richard et al. 2021; Ngan et al. 2023; Sakamuru et al. 2022). This approach should be further discussed by the scientific and regulatory community, to define priorities among the many imaging-based in vitro methods, identify funding sources, and avoid potential monopolies. Contract Research Organisations (CROs) promise to be a key player for implementing and making available imaging-based in vitro methods for regulatory use, but there are market- and business-related issues to be considered.

Quality assurance and data transparency

Existing GLP principles may not be fully applicable to studies based on emerging technologies, prompting various consortia to investigate different approaches to adapt quality assurance and system audit trails. A significant step forward was the publication of the Guidance Document on Good In Vitro Method Practices (GIVIMP) by OECD in 2018 (OECD 2018b). The document provides guidance for development and implementation of cellular methods generating data for regulatory safety assessment. Even if imaging technologies are not specifically covered, other important aspects related to in vitro methods (*e.g.,* facilities, test systems, reference/control items, SOPs) are explained. Complementarily, Good Cell Culture Practices (Pamies et al. 2022; Pistollato et al. 2022) were also developed by the research community to face challenges of advanced cellular models, including stem cells and organoids (Pamies et al. 2018), MicroPhysiological Systems and Organ-on-Chip (Pamies et al. 2024). GLP principles also apply to computerised systems handling complex algorithms and big datasets, as described in OECD documents, with similar principles also used in Good Clinical Trial practices.[9] Validation of traditional calculations, including statistics, often involve comparing the results obtained with the chosen methods (*e.g.,* Excel sheet files) to calculations performed either manually or with another system. GLP principles require long-term accessibility of the archived study data, including proof of its integrity. Inspectors require access to training datasets, software or source code, and need to understand how the data is managed, as well as have evidence of validated computerised systems. Importantly, GLP consider computerised system the combination of the software, the hardware, and the interface, making it necessary to validate again when changing one of these elements. Striking a balance between providing enough information for verification and keeping documentation efficient is therefore crucial. Last but not least, GLP principles require that the data is stored for a long period, up to 15 years in some jurisdictions. Storing vast amount of raw – and potentially processed – data for such extended period presents significant challenges in terms of cost and accessibility. Ensuring data integrity and future compatibility with evolving software and hardware is an added complication.

Algorithm complexity could be a challenge for data transparency. Different image segmentation techniques are available and factors like image complexity, desired level of detail, and computational resources should be taken

[6] https://globalbioimaging.org/

[7] https://quarep.org/

[8] https://www.eurobioimaging.eu/

[9] https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-computerised-systems-and-electronic-data-clinical-trials_en.pdf

into consideration when selecting the best approach (*e.g.,* thresholding, edge detection, region growing) (Shroff et al. 2024). Commercial software offers advantages such as tight integration with the microscope if provided directly by the microscopy vendors, ease to use of ready-made and well documented workflows and effective long-term support. The downside is the proprietary nature of the used algorithms with little insight of their precise makeup and lacking fully transparent validation and benchmarking against other solutions. Open-source solutions based on software platforms such as Fiji/ImageJ (Rueden et al. 2017), CellProfiler (Stirling et al. 2021; Soliman 2015) and pipelining software KNIME (Berthold et al. 2008) offer complete transparency of their code and are well suited to create bespoke workflows for specific use cases, which can become particularly powerful when combined with open-source data analysis based on R (Meyer et al. 2024; Pau et al. 2010) or Python (Rueden et al. 2022).

Recently DL, a specialised form of ML algorithms using Deep Neural Networks, is increasingly applied for microscopy image and data analysis (Moen et al. 2019; Seo et al. 2020). These algorithms, frequently summarised using the term AI, can be trained on big datasets, where each tissue or cell and its substructures are meticulously outlined by human experts. Through this training set, the algorithms learn to identify cell morphology, intensity variations, and subtle differences between cellular components. Once objects of interest are successfully identified, the analysis can proceed by extracting relevant features, including the size and shape of the identified objects, as well as the intensity and distribution of the fluorescent markers used in the experiment, or texture patterns within the objects. How data-driven models arrive at the predictions should be explained and the quality of unsupervised data analysis for pre-model training should be rigorously assessed. AI-based systems have the potential to mitigate challenges like diagnostic drift and observer variability. It is important to highlight, however, that while AI algorithms excel at automated tasks, human expertise remains irreplaceable for confirming and interpreting the outcome and toxicologists will maintain responsibility for final interpretation and decision-making. Verifying complex AI algorithms is challenging but can be achieved through rigorous auditable software engineering practices, including comprehensive documentation of development processes, robust risk-driven test suites, and source code reviews. Verification typically involves meticulous design of tests to assess the performance of different software units and functions at multiple levels. Explainable AI is a concept developed by the High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence (EU-HLEG-AI).[10] While the description of DL algorithms is unnecessary at this stage, establishing checkpoints throughout the process, such as validating training datasets, can ensure the validity of the data. Similar considerations also guided both FDA and EMA in the drafting of the "Good Machine Learning Practice for Medical Device Development: Guiding Principles" for use of AI/ML-based software in medical devices[11] and of the "Reflection paper on the use of AI in the medicinal product lifecycle",[12] respectively. Of note, the impact of AI on the long-established field of histopathology needs to be further defined when the data generated are used for regulatory purposes (Turner et al. 2020).

Advancing transparency and facilitating regulatory compliance also requires standardised data reporting practices, which should include the incorporation of tracking and version control for training data, models, and algorithms to ensure reproducibility of research findings. Even if open-source software is frequently used by the research community and is supposed to promote transparency, it may not be a compatible strategy for protecting companies' intellectual property and commercial interests. Standardisation of imaging metadata is essential for optimal data sharing, comparison, interpretation, and reusability across platforms. Metadata should be embedded within the image file itself or provided in supplementary files alongside the published images (Linkert et al. 2010; Hammer et al. 2021). However, current practices, often based on time-consuming manual curation in paper notebooks, are inefficient, inconsistent, and potentially error prone. Determining which metadata should be captured and the optimal timing and responsability for metadata entry during imaging workflow design are debated, with some advocating for real-time entry and others for post-acquisition entry. Consortia like Research Data Alliance (RDA),[13] GoFAIR,[14] FAIRsharing[15] and Open Microscopy Environment (OME)[16] are actively working to address standardisation issues for research data and trying to create solutions for image data management, including how data are organised, stored, and preserved. Using existing image data management systems like Open Microscopy Environment Remote Objects (OMERO) (Allan et al. 2012;

[10] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[11] https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

[12] https://www.ema.europa.eu/en/news/reflection-paper-use-artificial-intelligence-lifecycle-medicines

[13] https://www.rd-alliance.org/

[14] https://www.go-fair.org/

[15] https://fairsharing.org/

[16] https://www.openmicroscopy.org/

Burel et al. 2015; Li et al. 2016) and storing biological image datasets in repositories (*e.g.,* BioImage Archive,[17] IDR,[18] PIDAR[19]) can facilitate secure data sharing and foster collaboration (Williams et al. 2017; Ellenberg et al. 2018; Hartley et al. 2022). Overall, a collaborative data management approach focused on data security, transparency, traceability, and adherence to FAIR (Findability, Accessibility, Interoperability, and Reusability) principles is recommended (Weissgerber et al. 2024). To achieve these goals, data should be findable through registration in a searchable and publicly accessible on repository resource accompanied by rich metadata. Standardised and open communication protocols should be established to ensure data accessibility. Interoperability, facilitated by open, standardised formats and a common user-friendly metadata ontology, is needed to make sure that data can be processed by different software and hardware.

Data reliability and use in decision making

Assessing the reliability of the data generated with an imaging-based in vitro method is one of the main outcomes of validation studies. However, each data consumer has slightly different needs and approaches when assessing reliability of a test method for their own purposes. Method developers and industry experts need to understand all technicalities and instrument specifications to be able to implement a new test method and offer it to clients. When taking a decision concerning human safety, regulators need to have a clear understanding of the data interpretation procedure and of all possible uncertainties. For imaging-based in vitro methods, it is important to be able to decouple the uncertainty that is generated by the biological model itself from the one generated by the experimental procedure, including both the image acquisition and analysis processes. Currently users perform proficiency testing of the whole test method by analysing its performance on reference compounds. This evaluation relies on expert judgement by developers/end-users, and it is not clearly embedded in the test method definition, thus making it extremely difficult for the regulators to have a systematic way to measure uncertainties and assess data reliability.

The intended use of data (*e.g.,* regulatory submission, biomedical research) should determine the appropriate workflow and reporting standards, to guarantee that stakeholders have the complete set of information to perform a reliability assessment and use the data generated by the test method. Thus, data consumers should have a say in how the data is reported, including both metadata and ontology. In the field of omics, another emerging technology with the potential to widely impact regulatory sciences, the OECD is developing templates for data reporting (Harrill et al. 2021; OECD 2023b). This exercise, performed by international experts in a community-driven way, is ensuring that the data generated with omics technologies are reported in a harmonised way, thus allowing the regulators to clearly understand how the data are generated and how reliable they are, and therefore, to use them for decision-making. Using these types of templates, however, still leaves to the developers and end-users the freedom to choose their preferred instrument and technique to generate the data (Krebs et al. 2019). A similar approach in standardising data reporting could be proposed for imaging-based in vitro methods, to translate the use of these technologies across different sectors. Very comprehensive and clear reporting templates were already proposed in the literature (Nelson et al. 2021; Schmied et al. 2024; Sarkans et al. 2021; Montero Llopis et al. 2021) and they are an optimal starting point to develop reporting templates and guidelines that are specific for regulatory use and decision-making processes.

Implementation in CROs

CROs typically offer a variety of ready-to-go assays, including phenotypic, biochemical, cell-based, and toxicity assays, often combined with high-throughput screening capabilities. This positions CROs as key players in facilitating the adoption of imaging-based in vitro methods across both small companies and big pharma. Given the nature of their services, CROs are inherently structured to implementing rigorous SOPs and quality control measures. There is potential future demand for GLP-compliant imaging, but its development is hindered by limited solutions and slow user adoption. Currently, there is low demand for imaging instruments capable to fulfil all GLP requirements for computerised systems and CROs performing GLP studies rarely use imaging technologies for regulatory toxicology, with the notable exception of histopathology. Most of the CROs are mainly focusing on investigative cell-based imaging methods used upstream in the drug development phase, rather than in vitro assays for regulatory dossiers. However, driven by increased outsourcing from pharmaceutical and biotechnology companies, the market demand for cell-based imaging is experiencing significant growth and CROs are expanding their offerings to include these services for applications such as primary screening, target identification, compound profiling, and toxicity studies.

Several factors influence the selection of an imaging platform for integration into CROs business workflow. Key considerations include technical parameters like imaging resolution, instrument speed, multiplexing capabilities,

[17] https://www.ebi.ac.uk/bioimage-archive/

[18] https://idr.openmicroscopy.org/

[19] https://pidar.hpc4ai.unito.it/

automation level. User-friendliness of the software, robustness of the data generated, and workflow standardisation are equally important. A common approach is to opt for imaging systems that can automate image acquisition, analysis, archiving, and visualisation, providing end-to-end solutions. Manufacturer support and training are also important for the proper maintenance of the imaging instruments and software. Ideally, the imaging platform should be flexible and accommodate diverse types of assays such as live/fixed cell acquisition, 2D and 3D imaging, confocal capability, and both fluorescence and wide field imaging. For image and data analysis, commercial software is often the preferred choice since it provides standardised workflows that facilitate SOP development, while making a compromise with transparency. Data security is a top priority for CROs. To protect client information, they frequently utilise offline storage systems with dedicated servers, all located on-site, that ensure quick access and retrieval of data when needed for new analyses or revisiting previous analyses. Despite the potential, CROs face significant challenges in developing and utilising in vitro methods. High costs associated with HCI instruments, high-performance IT infrastructure, and GLP certification present a significant financial barrier. Finding personnel with the necessary multidisciplinary expertise in microscopy, biology imaging, and data analysis can be difficult. Balancing client demands for relevant testing with the financial risks of adopting new solutions is another obstacle, especially for smaller companies.

Due to their expertise and capabilities, CROs could serve as potential hubs for generating regulatory-compliant workflows. Transferability of methods to CROs is therefore crucial, but more public and private investment is needed to implement standardised imaging-based in vitro methods and ensure their availability to end-users. Developing and sharing standardised toxicology assays based on HCI and integrating them with omics approaches for a more comprehensive analysis represent exciting future business opportunities for CROs.

Conclusion

Facilitating the uptake of imaging-based in vitro methods for regulatory applications in biomedicine and toxicology necessitates further scientific and practical efforts across the whole experimental process of data generation, to increase confidence in the data used for decision-making. Systematic instrument calibration and shared performance assessments are essential for generating comparable test method outcomes across laboratories. The ideal image analysis workflow should foster innovation while maintaining consistent practices. Successful and trustworthy AI-powered image and data analysis should rely on transparent documentation of methods and code, defined workflows, as well as rigorous model validation. This includes accurate data annotation, robust validation sets, and thorough evaluation of accuracy and reproducibility with established quantitative metrics. Ideally, collaborative teams will work together on the strategy for imaging-based in vitro methods and drive standardisation initiatives for regulatory acceptance. These foundational elements will enable the incorporation of imaging methods in OECD TGs and their subsequent implementation under a quality management system across multiple CROs.

Acknowledgements The authors would like to thank Timothy Bolus, Peter Macko, Roland Nitschke, Christian Strueh and Özlem Yavas for the insightful discussions.

Declarations

Conflict of interest Erio Barale-Thomas, David Egan, Katharina Koch, Elena Miranda, Enrico Pesenti, Fernanda Ricci, Alexander Schreiner, Nadine Stokar-Regenscheit, Fredrik C. Wieland and Amy Wilson are employed by the respective companies. Elena Miranda holds shares in GSK. Amy Wilson holds shares in AstraZeneca. Nadine Stokar-Regenscheit holds share in F.Hoffmann-La Roche Ldt. Katharina Koch holds share in DNTOX. Philipp Kainz is a director and shareholder of KOLAIDO GmbH. David Egan is a director and shareholder of Core Life Analytics BV. Iris Mangas contributed in particular to the section "Developmental neurotoxicity in vitro battery" of this publication. The views expressed, and arguments employed in this article are those of the authors and do not necessarily reflect the official views or policies of their companies or institutions of employment or of the governments of its member countries.

References

Allan C, Burel J-M, Moore J, Blackburn C, Linkert M, Loynton S, MacDonald D et al (2012) OMERO: flexible, model-driven data management for experimental biology. Nat Methods 9(3):245–253. https://doi.org/10.1038/nmeth.1896

Arevalo J, Ellen S, van Dijk R, Carpenter AE, Singh S (2024) Evaluating batch correction methods for image-based cell profiling. bioRxiv. https://doi.org/10.1101/2023.09.15.558001

Bal-Price A, Hogberg HT, Crofton KM, Daneshian M, FitzGerald RE, Fritsche E, Heinonen T et al (2018) Recommendation on test readiness criteria for new approach methods in toxicology: exemplified for developmental neurotoxicity. ALTEX 35(3):306–352. https://doi.org/10.14573/altex.1712081

Barisoni L, Lafata KJ, Hewitt SM, Madabhushi A, Balis UGJ (2020) Digital pathology and computational image analysis in nephropathology. Nat Rev Nephrol 16(11):669–685. https://doi.org/10.1038/s41581-020-0321-6

Berthold MR, Nicolas C, Fabian D, Thomas RG, Tobias K, Thorsten M, Peter O, Christoph S, Kilian T, Bernd W (2008) KNIME: The Konstanz Information Miner. In: Christine P, Hans B, Lars ST, Reinhold D (eds) Data analysis, machine learning and applications. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 319–26

Blum J, Masjosthusmann S, Bartmann K, Bendt F, Dolde X, Dönmez A, Förster N et al (2023) Establishment of a human cell-based in vitro battery to assess developmental neurotoxicity hazard of chemicals. Chemosphere 311:137035. https://doi.org/10.1016/j.chemosphere.2022.137035

Boisclair J, Bawa B, Barale-Thomas E, Bertrand L, Carter J, Crossland R, Dorn C et al (2022) IT/QA and regulatory aspects of digital pathology: results of the 8th ESTP international workshop. Toxicol Pathol 50(6):793–807. https://doi.org/10.1177/01926233221113275

Bray M-A, Singh S, Han H, Davis CT, Borgeson B, Hartland C, Kost-Alimova M, Gustafsdottir SM, Gibson CC, Carpenter AE (2016) Cell painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. Nat Protoc 11(9):1757–1774. https://doi.org/10.1038/nprot.2016.105

Burel J-M, Besson S, Blackburn C, Carroll M, Ferguson RK, Flynn H, Gillen K et al (2015) Publishing and sharing multi-dimensional image data with OMERO. Mamm Genome 26(9):441–447. https://doi.org/10.1007/s00335-015-9587-6

Caicedo JC, Cooper S, Heigwer F, Warchal S, Qiu P, Molnar C, Vasilevich AS et al (2017) Data-analysis strategies for image-based cell profiling. Nat Methods 14(9):849–863. https://doi.org/10.1038/nmeth.4397

Chandrasekaran SN, Ceulemans H, Boyd JD, Carpenter AE (2021) Image-based profiling for drug discovery: due for a machine-learning upgrade? Nat Rev Drug Discov 20(2):145–159. https://doi.org/10.1038/s41573-020-00117-w

Chandrasekaran SN, Cimini BA, Goodale A, Miller L, Kost-Alimova M, Jamali N, Doench JG et al (2024) Three million images and morphological profiles of cells treated with matched chemical and genetic perturbations. Nat Methods 21(6):1114–1121. https://doi.org/10.1038/s41592-024-02241-6

Cimini BA, Chandrasekaran SN, Kost-Alimova M, Miller L, Goodale A, Fritchman B, Byrne P et al (2023) Optimizing the cell painting assay for image-based profiling. Nat Protoc 18(7):1981–2013. https://doi.org/10.1038/s41596-023-00840-9

Crofton KM, Mundy WR, Lein PJ, Bal-Price A, Coecke S, Seiler AEM, Knaut H, Buzanska L, Goldberg A (2011) Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals. Altern Anim Exp 28(1):9–15. https://doi.org/10.14573/altex.2011.1.009

Dal Negro G, Eskes C, Belz S, Bertein C, Chlebus M, Corvaro M, Corvi R et al (2018) One science-driven approach for the regulatory implementation of alternative methods: a multi-sector perspective. Regul Toxicol Pharmacol 99:33–49. https://doi.org/10.1016/j.yrtph.2018.08.002

EFSA, Antonio Hernandez-Jerez, Tamara Coja, Martin Paparella, Anna Price, Jerome Henri, Andreas Focks, et al. (2024) Statement on the toxicological properties and maximum residue levels of acetamiprid and its metabolites. EFSA J 22(5):e8759. https://doi.org/10.2903/j.efsa.2024.8759

Ellenberg J, Swedlow JR, Barlow M, Cook CE, Sarkans U, Patwardhan A, Brazma A, Birney E (2018) A call for public archives for biological image data. Nat Methods 15(11):849–854. https://doi.org/10.1038/s41592-018-0195-8

FDA (2016) Technical performance assessment of digital pathology whole slide imaging devices - guidance for industry and food and drug administration staff

FDA (2023) Use of whole slide imaging in nonclinical toxicology studies: questions and answers - guidance for industry. https://www.fda.gov/media/168431/download

U.S. Food and Drug Administration (2018) Clinical trial imaging endpoint process standards - guidance for industry. https://www.fda.gov/media/81172/download.

Förster N, Butke J, Keßel HE, Bendt F, Melanie Pahl Lu, Li XF et al (2022) Reliable identification and quantification of neural cells in microscopic images of neurospheres. Cytometry A 101(5):411–422. https://doi.org/10.1002/cyto.a.24514

Gustafsdottir SM, Ljosa V, Sokolnicki KL, Anthony Wilson J, Walpita D, Kemp MM, Seiler KP et al (2013) Multiplex cytological profiling assay to measure diverse cellular states. PLoS ONE 8(12):e80999. https://doi.org/10.1371/journal.pone.0080999

Halter M, Bier E, DeRose PC, Cooksey GA, Choquette SJ, Plant AL, Elliott JT (2014) An automated protocol for performance benchmarking a widefield fluorescence microscope. Cytometry Part A 85(11):978–85. https://doi.org/10.1002/cyto.a.22519

Hammer M, Huisman M, Rigano A, Boehm U, Chambers JJ, Gaudreault N, North AJ et al (2021) Towards community-driven metadata standards for light microscopy: tiered specifications extending the OME model. Nat Methods 18(12):1427–1440. https://doi.org/10.1038/s41592-021-01327-9

Harrill JA, Freudenrich T, Wallace K, Ball K, Shafer TJ, Mundy WR (2018) Testing for developmental neurotoxicity using a battery of in vitro assays for key cellular events in neurodevelopment. Toxicol Appl Pharmacol 354:24–39. https://doi.org/10.1016/j.taap.2018.04.001

Harrill JA, Viant MR, Yauk CL, Sachana M, Gant TW, Auerbach SS, Beger RD et al (2021) Progress towards an OECD reporting framework for transcriptomics and metabolomics in regulatory toxicology. Regul Toxicol Pharmacol 125:105020. https://doi.org/10.1016/j.yrtph.2021.105020

Hartley M, Kleywegt GJ, Patwardhan A, Sarkans U, Swedlow JR, Brazma A (2022) The bioimage archive – building a home for life-sciences microscopy data. Comput Resour Mol Biol 434(11):167505. https://doi.org/10.1016/j.jmb.2022.167505

Hosseini R, Vlasveld M, Willemse J, van de Water B, Le Dévédec SE, Wolstencroft KJ (2023) FAIR high content screening in bioimaging. Sci Data 10(1):462. https://doi.org/10.1038/s41597-023-02367-w

Huisman, Maximiliaan, Mathias Hammer, Alex Rigano, Ulrike Boehm, James J. Chambers, Nathalie Gaudreault, Alison J. North, et al. (2021) A perspective on microscopy metadata: data provenance and quality control. https://arxiv.org/abs/1910.11370

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) (2023a) ICH Q2(R2) - validation of analytical procedures. https://database.ich.org/sites/default/files/ICH_Q2%28R2%29_Guideline_2023_1130.pdf

International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) (2023b) ICH Q14 - analytical procedure development. https://database.ich.org/sites/default/files/ICH_Q14_Guideline_2023_1116.pdf

Jacobsen M, Lewis A, Baily J, Fraser A, Rudmann D, Ryan S (2021) Utilizing whole slide images for the primary evaluation and peer review of a GLP-compliant rodent toxicology study. Toxicol Pathol 49(6):1164–1173. https://doi.org/10.1177/0192623321017031

Klutzny S, Kornhuber M, Morger A, Schönfelder G, Volkamer A, Oelgeschläger M, Dunst S (2022) Quantitative high-throughput phenotypic screening for environmental estrogens using the e-morph screening assay in combination with in silico predictions.

Environ Int 158:106947. https://doi.org/10.1016/j.envint.2021.106947

Koch K, Bartmann K, Hartmann J, Kapr J, Klose J, Kuchovská E, Pahl M, Schlüppmann K, Zühr E, Fritsche E (2022) Scientific validation of human neurosphere assays for developmental neurotoxicity evaluation. Front Toxicol 4:816370. https://doi.org/10.3389/ftox.2022.816370

Kolle SN, Moreno MCR, Mayer W, van Cott A, van Ravenzwaay B, Landsiedel R (2015) The EpiOcular™ eye irritation test is the method of choice for the in vitro eye irritation testing of agrochemical formulations: correlation analysis of epiocular eye irritation test and BCOP test data according to the UN GHS, US EPA and Brazil ANVISA classification schemes. Altern Lab Anim 43(3):181–198. https://doi.org/10.1177/0261192915043000307

Krebs A, Waldmann T, Wilks MF, Van Vugt-Lussenburg BMA, Van der Burg B, Terron A, Steger-Hartmann T et al (2019) Template for the description of cell-based toxicological test methods to allow evaluation and regulatory use of the data. ALTEX 36(4):682–99. https://doi.org/10.14573/altex.1909271

Kunis S, Hänsch S, Schmidt C, Wong F, Strambio-De-Castillia C, Weidtkamp-Peters S (2021) MDEmic: a metadata annotation tool to facilitate management of FAIR image data in the bioimaging community. Nat Methods 18(12):1416–1417. https://doi.org/10.1038/s41592-021-01288-z

Kwee E, Peterson A, Halter M, Elliott J (2021) Practical application of microsphere samples for benchmarking a quantitative phase imaging system. Cytometry Part A 99(10):1022–32. https://doi.org/10.1002/cyto.a.24291

Laine RF, Arganda-Carreras I, Henriques R, Jacquemet G (2021) Avoiding a replication crisis in deep-learning-based bioimage analysis. Nat Methods 18(10):1136–1144. https://doi.org/10.1038/s41592-021-01284-3

Li S, Besson S, Blackburn C, Carroll M, Ferguson RK, Flynn H, Gillen K et al (2016) Metadata management for high content screening in OMERO. Methods 96:27–32. https://doi.org/10.1016/j.ymeth.2015.10.006

Linkert M, Rueden CT, Allan C, Burel J-M, Moore W, Patterson A, Loranger B et al (2010) Metadata matters: access to image data in the real world. J Cell Biol 189(5):777–782. https://doi.org/10.1083/jcb.201004104

Llopis M, Paula RA, Senft TJ, Ross-Elliott RS, Keeley DP, Koshar P, Marqués G et al (2021) Best practices and tools for reporting reproducible fluorescence microscopy methods. Nat Methods 18(12):1463–1476. https://doi.org/10.1038/s41592-021-01156-w

Maier-Hein L, Reinke A, Godau P, Tizabi MD, Buettner F, Christodoulou E, Glocker B et al (2024) Metrics reloaded: recommendations for image analysis validation. Nat Methods 21(2):195–212. https://doi.org/10.1038/s41592-023-02151-z

Martin MM, Baker NC, Boyes WK, Carstens KE, Culbreth ME, Gilbert ME, Harrill JA et al (2022) An expert-driven literature review of "negative" chemicals for developmental neurotoxicity (DNT) in vitro assay evaluation. Neurotoxicol Teratol 93:107117. https://doi.org/10.1016/j.ntt.2022.107117

Marx V (2022) The making of microscope camera standards. Nat Methods 19(7):788–791. https://doi.org/10.1038/s41592-022-01545-9

Masjosthusmann S, Blum J, Bartmann K, Dolde X, Holzer A-K, Stürzl L-C, Keßel EH et al (2020) Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. EFSA Support Publ 17(10):1938E. https://doi.org/10.2903/sp.efsa.2020.EN-1938

McGenity C, Clarke EL, Jennings C, Matthews G, Cartlidge C, Freduah-Agyemang H, Stocken DD, Treanor D (2024) Artificial intelligence in digital pathology: a systematic review and meta-analysis of diagnostic test accuracy. NPJ Digit Med 7(1):114. https://doi.org/10.1038/s41746-024-01106-8

Mehrvar S, Himmel LE, Babburi P, Goldberg AL, Guffroy M, Janardhan K, Krempley AL, Bawa B (2021) Deep learning approaches and applications in toxicologic histopathology: current status and future perspectives. J Pathol Inf 12:42. https://doi.org/10.4103/jpi.jpi_36_21

Meyer L, Eling N, Bodenmiller B (2024) Cytoviewer: an R/Bioconductor package for interactive visualization and exploration of highly multiplexed imaging data. BMC Bioinform 25(1):9. https://doi.org/10.1186/s12859-023-05546-z

Miura K, Nørrelykke SF (2021) reproducible image handling and analysis. EMBO J 40(3):e105889. https://doi.org/10.15252/embj.2020105889

Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D (2019) Deep learning for cellular image analysis. Nat Methods 16(12):1233–1246. https://doi.org/10.1038/s41592-019-0403-1

Moore J, Allan C, Besson S, Burel J-M, Diel E, Gault D, Kozlowski K et al (2021) OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies. Nat Methods 18(12):1496–1498. https://doi.org/10.1038/s41592-021-01326-w

Moulin P, Grünberg K, Barale-Thomas E, van der Laak J (2021) IMI-Bigpicture: a central repository for digital pathology. Toxicol Pathol 49(4):711–713. https://doi.org/10.1177/0192623321989644

Mundy WR, Padilla S, Breier JM, Crofton KM, Gilbert ME, Herr DW, Jensen KF et al (2015) Expanding the test set: chemicals with potential to disrupt mammalian brain development. Neurotoxicol Teratol 52:25–35. https://doi.org/10.1016/j.ntt.2015.10.001

National Academies of Sciences, Engineering and Medicine (2023) Building confidence in new evidence streams for human health risk assessment: lessons learned from laboratory mammalian toxicity tests. The National Academies Press, Washington, DC

Nelson G, Boehm U, Bagley S, Bajcsy P, Bischof J, Brown CM, Dauphin A et al (2021) QUAREP-LiMi: a community-driven initiative to establish guidelines for quality assessment and reproducibility for instruments and images in light microscopy. J Microsc 284(1):56–73. https://doi.org/10.1111/jmi.13041

Ngan DK, Xia M, Simeonov A, Huang R (2023) In vitro profiling of pesticides within the Tox21 10K compound library for bioactivity and potential toxicity. Toxicol Appl Pharmacol 473:116600. https://doi.org/10.1016/j.taap.2023.116600

Nyffeler J, Willis C, Lougee R, Richard A, Paul-Friedman K, Harrill JA (2020) Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling. Toxicol Appl Pharmacol 389:114876. https://doi.org/10.1016/j.taap.2019.114876

Nyffeler J, Willis C, Harris FR, Foster MJ, Chambers B, Culbreth M, Brockway RE et al (2023) Application of cell painting for chemical hazard evaluation in support of screening-level chemical assessments. Toxicol Appl Pharmacol 468:116513. https://doi.org/10.1016/j.taap.2023.116513

Odje F, Meijer D, Von Coburg E, van der Hooft JJJ, Dunst S, Medema MH, Volkamer A (2024) Unleashing the potential of cell painting assays for compound activities and hazards prediction. Front Toxicol. https://doi.org/10.3389/ftox.2024.1401036

OECD (2005) Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment - series on testing and assessment No. 34. OECD Publishing, Paris. https://one.oecd.org/document/ENV/JM/MONO(2005)14/en/pdf

OECD (2007) Test No. 426: developmental neurotoxicity study. OECD Publishing, Paris

OECD (2018) Guidance document on good in vitro method practices (GIVIMP) - series on testing and assessment No. 286. OECD Publishing, Paris

OECD (2018a) Extended one-generation reproductive toxicity study (EOGRTS) (OECD TG 443). https://www.oecd-ilibrary.org/content/component/9789264304741-34-en

OECD (2023a) Initial recommendations on evaluation of data from the developmental neurotoxicity (DNT) in-vitro testing battery - series on testing & assessment No. 377. OECD Publishing, Paris. https://one.oecd.org/document/ENV/CBC/MONO(2023)13/en/pdf.

OECD (2023b) OECD omics reporting framework (OORF): Guidance on reporting elements for the regulatory use of omics data from laboratory-based toxicology studies - series on testing and assessment No. 390. OECD Publishing, Paris. https://one.oecd.org/document/ENV/CBC/MONO(2023)41/en/pdf

Pamies D, Bal-Price A, Chesné C, Coecke S, Dinnyes A, Eskes C, Grillari R et al (2018) Advanced good cell culture practice for human primary, stem cell-derived and organoid models as well as microphysiological systems. ALTEX 35(3):353–78. https://doi.org/10.14573/altex.1710081

Pamies D, Leist M, Coecke S, Bowe G, Allen DG, Gstraunthaler G, Bal-Price Anna et al (2022) Guidance document on good cell and tissue culture practice 2.0 (GCCP 2.0). ALTEX 39:30–70. https://doi.org/10.14573/altex.2111011

Pamies D, Ekert J, Zurich M-G, Frey O, Werner S, Piergiovanni M, Freedman BS et al (2024) Recommendations on fit-for-purpose criteria to establish quality management for microphysiological systems and for monitoring their reproducibility. Stem Cell Rep 19(5):604–617. https://doi.org/10.1016/j.stemcr.2024.03.009

Pau G, Fuchs F, Sklyar O, Boutros M, Huber W (2010) EBImage–an R package for image processing with applications to cellular phenotypes. Bioinform Oxford Engl 26(7):979–981. https://doi.org/10.1093/bioinformatics/btq046

Pistollato F, Madia F, Corvi R, Munn S, Grignard E, Paini A, Worth A et al (2021) Current EU regulatory requirements for the assessment of chemicals and cosmetic products: challenges and opportunities for introducing new approach methodologies. Arch Toxicol 95(6):1867–1897. https://doi.org/10.1007/s00204-021-03034-y

Pistollato F, Bal-Price A, Coecke S, Parvatam S, Pamies D, Czysz K, Hao J et al (2022) Quality criteria for in vitro human pluripotent stem cell-derived models of tissue-based cells. Reprod Toxicol Elmsford NY 112:36–50. https://doi.org/10.1016/j.reprotox.2022.06.003

Richard AM, Huang R, Waidyanatha S, Shinn P, Collins BJ, Thillainadarajah I, Grulke CM et al (2021) The Tox21 10K compound library: collaborative chemistry advancing toxicology. Chem Res Toxicol 34(2):189–216. https://doi.org/10.1021/acs.chemrestox.0c00264

Rigano A, Ehmsen S, Öztürk SU, Ryan J, Balashov A, Hammer M, Kirli K et al (2021) Micro-meta app: an interactive tool for collecting microscopy metadata based on community specifications. Nat Methods 18(12):1489–1495. https://doi.org/10.1038/s41592-021-01315-z

Rivenson Y, Wang H, Wei Z, de Haan K, Zhang Y, Yichen Wu, Günaydın H et al (2019) Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. Nat Biomed Eng 3(6):466–477. https://doi.org/10.1038/s41551-019-0362-y

Rudmann DG, Bertrand L, Zuraw A, Deiters J, Staup M, Rivenson Y, Kuklyte J (2023) Building a nonclinical pathology laboratory of the future for pharmaceutical research excellence. Drug Discov Today 28(10):103747. https://doi.org/10.1016/j.drudis.2023.103747

Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, Eliceiri KW (2017) ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinform 18(1):529. https://doi.org/10.1186/s12859-017-1934-z

Rueden CT, Hiner MC, Evans EL, Pinkert MA, Lucas AM, Carpenter AE, Cimini BA, Eliceiri KW (2022) PyImageJ: a library for integrating ImageJ and python. Nat Methods 19(11):1326–1327. https://doi.org/10.1038/s41592-022-01655-4

Russell WMS, Rex LB, Charles WH (1959) The principles of humane experimental technique, vol 238. Methuen, London

Ryan J, Pengo T, Rigano A, Llopis PM, Itano MS, Cameron LA, Marqués G, Strambio-De-Castillia C, Sanders MA, Brown CM (2021) MethodsJ2: a software tool to capture metadata and generate comprehensive microscopy methods text. Nat Methods 18(12):1414–1416. https://doi.org/10.1038/s41592-021-01290-5

Sachana M, Shafer TJ, Terron A (2021) Toward a better testing paradigm for developmental neurotoxicity: OECD efforts and regulatory considerations. Biology. https://doi.org/10.3390/biology10020086

Sakamuru S, Huang R, Xia M (2022) Use of Tox21 screening data to evaluate the COVID-19 drug candidates for their potential toxic effects and related pathways. Front Pharmacol 13:935399. https://doi.org/10.3389/fphar.2022.935399

Sarkans U, Chiu W, Collinson L, Darrow MC, Ellenberg J, Grunwald D, Hériché J-K et al (2021) REMBI: recommended metadata for biological images—enabling reuse of microscopy data in biology. Nat Methods 18(12):1418–1422. https://doi.org/10.1038/s41592-021-01166-8

Schmidt C, Hanne J, Moore J, Meesters C, Ferrando-May E, Weidtkamp-Peters S (2022) Research data management for bioimaging: the 2021 NFDI4BIOIMAGE community survey. F1000Research 11:638. https://doi.org/10.12688/f1000research.121714.2

Schmied C, Nelson MS, Avilov S, Bakker G-J, Bertocchi C, Bischof J, Boehm U et al (2024) Community-developed checklists for publishing images and image analyses. Nat Methods 21(2):170–181. https://doi.org/10.1038/s41592-023-01987-9

Schmuck MR, Temme T, Dach K, de Boer D, Barenys M, Bendt F, Mosig A, Fritsche E (2017) Omnisphero: a high-content image analysis (HCA) approach for phenotypic developmental neurotoxicity (DNT) screenings of organoid neurosphere cultures in vitro. Arch Toxicol 91(4):2017–2028. https://doi.org/10.1007/s00204-016-1852-2

Schwenk M, Werner M, Younes M (2002) Regulatory toxicology: objectives and tasks defined by the working group of the German society of experimental and clinical pharmacology and toxicology. Toxicol Lett 126(3):145–153. https://doi.org/10.1016/s0378-4274(01)00457-x

Senft RA, Diaz-Rohrer B, Colarusso P, Swift L, Jamali N, Jambor H, Pengo T et al (2023) A biologist's guide to planning and performing quantitative bioimaging experiments. PLoS Biol 21(6):1–22. https://doi.org/10.1371/journal.pbio.3002167

Seo H, Khuzani MB, Vasudevan V, Huang C, Ren H, Xiao R, Jia X, Xing L (2020) Machine learning techniques for biomedical image segmentation: an overview of technical aspects and introduction to state-of-art applications. Med Phys 47(5):e148–e167. https://doi.org/10.1002/mp.13649

Serafini MM, Sepehri S, Midali M, Stinckens M, Biesiekierska M, Wolniakowska A, Gatzios A et al (2024) Recent advances and current challenges of new approach methodologies in developmental and adult neurotoxicity testing. Arch Toxicol 98(5):1271–1295. https://doi.org/10.1007/s00204-024-03703-8

Shroff H, Testa I, Jug F, Manley S (2024) Live-cell imaging powered by computation. Nat Rev Mol Cell Biol 25(6):443–463. https://doi.org/10.1038/s41580-024-00702-6

Soliman K (2015) CellProfiler: novel automated image segmentation procedure for super-resolution microscopy. Biol Proced Online 17(1):11. https://doi.org/10.1186/s12575-015-0023-9

Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A (2021) CellProfiler 4: improvements in speed, utility and usability. BMC Bioinform 22(1):433. https://doi.org/10.1186/s12859-021-04344-9

Stokar-Regenscheit N, Bell L, Berridge B, Rudmann D, Tagle D, Hargrove-Grimes P, Schaudien D et al (2024) Complex in vitro model characterization for context of use in toxicologic pathology: use cases by collaborative teams of biologists, bioengineers, and pathologists. Toxicol Pathol. https://doi.org/10.1177/01926233241253811

Tromans-Coia C, Jamali N, Abbasi HS, Giuliano KA, Hagimoto M, Jan K, Kaneko E et al (2023) Assessing the performance of the cell painting assay across different imaging systems. Cytometry A 103(11):915–926. https://doi.org/10.1002/cyto.a.24786

Turner OC, Aeffner F, Bangari DS, High W, Knight B, Forest T, Cossic B et al (2020) Society of toxicologic pathology digital pathology and image analysis special interest group article*: opinion on the application of artificial intelligence and machine learning to digital toxicologic pathology. Toxicol Pathol 48(2):277–294. https://doi.org/10.1177/0192623319881401

Weissgerber TL, Gazda MA, Nilsonne G, ter Riet G, Cobey KD, Prieß-Buchheit J, Noro J et al (2024) Understanding the provenance and quality of methods is essential for responsible reuse of FAIR data. Nat Med 30(5):1220–1221. https://doi.org/10.1038/s41591-024-02879-x

Williams E, Moore J, Li SW, Rustici G, Tarkowska A, Chessel A, Leo S et al (2017) Image data resource: a bioimage data integration and publication platform. Nat Methods 14(8):775–781. https://doi.org/10.1038/nmeth.4326

Wolff C, Neuenschwander M, Beese CJ, Sitani D, Ramos MC, Srovnalova A, Varela MJ et al (2024) Morphological profiling dataset of EU- OPENSCREEN bioactive compounds over multiple imaging sites and cell lines. bioRxiv. https://doi.org/10.1101/2024.08.27.609964

Zuraw A, Aeffner F (2022) Whole-slide imaging, tissue image analysis, and artificial intelligence in veterinary pathology: an updated introduction and review. Vet Pathol 59(1):6–25. https://doi.org/10.1177/03009858211040484

Authors and Affiliations

Monica Piergiovanni[1] · Milena Mennecozzi[1] · Erio Barale-Thomas[2] · Davide Danovi[3] · Sebastian Dunst[4] · David Egan[5] · Aurora Fassi[1] · Matthew Hartley[6] · Philipp Kainz[22] · Katharina Koch[7,8] · Sylvia E. Le Dévédec[9] · Iris Mangas[10] · Elena Miranda[11] · Jo Nyffeler[12] · Enrico Pesenti[13] · Fernanda Ricci[14] · Christopher Schmied[15] · Alexander Schreiner[16] · Nadine Stokar-Regenscheit[17] · Jason R. Swedlow[18] · Virginie Uhlmann[19] · Fredrik C. Wieland[20] · Amy Wilson[21] · Maurice Whelan[1]

✉ Monica Piergiovanni
monica.piergiovanni@ec.europa.eu

[1] European Commission, Joint Research Centre (JRC), Ispra, Italy

[2] Preclinical Sciences and Translational Safety, Janssen Pharmaceuticals, Beerse, Belgium

[3] Department of Basic and Clinical Neuroscience, King's College London, London, UK

[4] German Centre for the Protection of Laboratory Animals (Bf3R), Department Experimental Toxicology and ZEBET, German Federal Institute for Risk Assessment, Berlin, Germany

[5] Core Life Analytics BV, 57 Kabelweg, 1014 BA Amsterdam, The Netherlands

[6] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, UK

[7] IUF – Leibniz Research Institute for Environmental Medicine, Duesseldorf, Germany

[8] DNTOX GmbH, Duesseldorf, Germany

[9] Leiden Academic Centre for Drug Research (LACDR), Faculty of Science, Leiden University, 2333 Leiden, Netherlands

[10] European Food Safety Authority (EFSA), Parma, Italy

[11] BioImaging, GSK, Stevenage SG1 2NY, UK

[12] Department of Ecotoxicology, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

[13] Crown Bioscience Inc, 16550 West Bernardo Drive, Building 5, Suite 525, San Diego, CA 92127, USA

[14] Axxam SpA, Bresso, Italy

[15] EU-OPENSCREEN ERIC, Campus Berlin-Buch, Robert-Roessle-Str. 10, 13125 Berlin, Germany

[16] Revvity Health Sciences, Inc, Waltham, MA, USA

[17] Roche Pharma Research and Early Development (pRED), Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland

[18] Divisions of Computational Biology and Molecular, Cell and Developmental Biology, School of Life Sciences, National Phenotypic Screening Centre, University of Dundee, Dundee, UK

[19] BioVisionCenter, University of Zürich, Zurich, Switzerland

[20] Life Science Business Europe, Yokogawa Deutschland GmbH, Ratingen, Germany

[21] Safety Sciences, Clinical Pharmacology and Safety Sciences, AstraZeneca, Cambridge, UK

[22] KOLAIDO GmbH, Altenrhein, Switzerland