Concept Article

The Benefits of Validation of Methods for Toxicity Testing Outweigh Its Costs

Anne Gourmelon[1], Philippe Hubert[2], Elise Grignard[2], Lisa Baumann[3], Sharon Munn[4] and Cécile Michel-Caillet[5]

[1]OECD Environment Directorate, Paris, France; [2]PEPPER, Paris, France; [3]Vrije Universiteit Amsterdam, Amsterdam Institute for Life and Environment, Section Environmental Health and Toxicology, Amsterdam, The Netherlands; [4]European Commission, Joint Research Centre (JRC), Ispra, Italy; [5]REACh-CLP-Endocrine Disrupters Unit, Risk Assessment Department, French Agency for Food, Environmental and Occupational Health Safety (ANSES), Maisons-Alfort, France

Received March 3, 2024;
Accepted March 13, 2024;
Epub March 19, 2024;

Correspondence:
Anne Gourmelon, PhD,
OECD Environment Directorate,
2 rue André-Pascal,
75775 Paris cedex 16, France
(Anne.Gourmelon@oecd.org)

ALTEX 41(3), 395-401.
doi:10.14573/altex.2403051

Abstract

The 4th Annual Forum on Endocrine Disrupters organized by the European Commission brought together the authors of this article around the topic: "From bench to validated test guidelines: (pre)validation of test methods". Validation activities are meant to demonstrate the relevance and reliability of methods and approaches used in regulatory safety testing. These activities are essential to facilitate regulatory use, still they are largely underfunded and unattractive to the scientific community. In the last decade, large amounts of funding have been invested in European research towards the development of approaches that can be used in regulatory decision-making, including for the identification of endocrine disrupters. There is a vast pool of candidate test methods for potential regulatory applications, but most of them will not be used due to the absence of consideration of their relevance and reliability outside the method developer's laboratory. This article explains the reasons why such a gap exists between the outputs of research projects and the uptake in a regulatory context. In parallel, there are also increasing expectations from the regulatory science community that validation becomes more efficient with respect to time and resources. This article shares some of the lessons learned and proposes paths forward for validation of new methods that are not intended as one-to-one replacements of animal studies. This includes submitting only mature methods for validation that were developed following good practices and good documentation, proposing a greater emphasis on well-documented transferability studies, and adopting a cost-sharing model among those who benefit from validated methods.

Plain language summary

Validation activities for methods intended to be used to assess chemical safety have a cost but also bring substantial benefits when the validated methods are established as OECD Test Guidelines, which results in mutual acceptance of data generated by the methods across OECD member and adhering countries. The article discusses some of the challenges faced when method validation is underfunded and unattractive for researchers. Proposals are made to improve the current situation, gain efficiency, and make validation a shared responsibility.

1 Why is validation needed?

Validation activities are essential in many areas where methodologies are developed to address specific questions. Chemical safety testing is one of these areas where methods are needed to assess the toxicity of chemicals. In this field, the methodologies currently rely primarily on testing in animal models, yet great efforts are being made to place less reliance on *in vivo* testing through the use of cell- or tissue- or *in chemico*-based test systems and/or use of predictive computational models. The number of health effects to investigate (target organ toxicity, reproductive toxicity, immunotoxicity, developmental immunotoxicity and neurotoxicity, carcinogenicity, endocrine disruption, etc.) together with the progress in the life sciences, bioengineering/biotechnologies, and computational capacity allow constant evolution and improvement of methods and approaches for better health and environmental protection.

Validation builds trust and confidence that methods used are relevant and reproducible, wherever and by whomever they are implemented. For the last few years, the regulatory science community has been reflecting on validation practices and opportunities of gaining efficiencies in the time and resources involved in validation activities (Piersma et al., 2018; Burgdorf et al., 2019; van der Zalm et al., 2022).

The authors of this paper consider that resources currently invested in validation are largely insufficient, and this is one of the factors limiting the pace of broad application of new approach methods (NAMs). While attention should be raised on the need for

more resources, stakeholders and the broader regulatory community are also interested in ways to optimize processes. Authors of this article participated in a dedicated session on validation at the 4th Forum on Endocrine Disrupters (ED Forum) organized by the European Commission in September 2022 (EU, 2022).

Whilst acknowledging that there are many organizations (e.g., ISO, CEN, ICH) involved in validation activities, here, we provide a focus on a few practical aspects of method validation in the context of OECD Test Guideline development, not pre-empting the work that will be performed by an expert group recently nominated to revise the OECD Guidance Document on validation and international acceptance (OECD, 2005). We attempt to provide an overview of resources, efforts, and preparatory work needed to carry out method validation in the context of OECD Test Guideline development, drawing from recent experiences at the national or global level to transform methods into validated methods. We also attempt to suggest possible paths towards more efficient validation.

2 Principles of validation

Validation is about demonstrating relevance and reproducibility. In the context of chemical hazard and risk assessment, relevance means biological relevance of the measured endpoint(s) in the assay to the entire organ system or organism level and its normal functioning. Concretely, it means that the method needs to measure a biologically relevant effect following exposure to a chemical that is active on the biological target, and the absence of effect following exposure to an inactive chemical.

Reproducibility means that results obtained are reproducible over time when experiments are repeated and that the method can be transferred successfully to naïve laboratories. A survey on the credibility crisis in scientific studies indicated that up to 60% of experiments cannot be reproduced within the same laboratory (Baker, 2016); this figure rises to 80% when other labs try to reproduce the results. Different factors were identified as causes for this reproducibility problem, among them the high pressure to publish scientific results ("publish or perish") and the failure to adhere to good scientific practices seem to be the major drivers. Good scientific practice in the field of toxicity testing is a key characteristic of validation that ensures a method is reproducible and thus ready for use in a regulatory context.

Validation forces the method developer: 1) to describe the protocol in sufficient detail to allow others to implement the method and reproduce the same results, 2) to define and explain the data interpretation procedure, and 3) to consider how much variability can be allowed without jeopardizing the outcome of the assay.

Demonstration of reproducibility is a key point in increasing societal confidence in science and regulation. These aspects are also foundational to the development of OECD Test Guidelines, which, together with Good Laboratory Practices, form the basis of the system of Mutual Acceptance of Data[1].

There are different ways to demonstrate a method is valid: (i) prospective validation (laboratory work); (ii) performance-based validation (against pre-existing standards); and (iii) retrospective validation (collecting and evaluating existing data that may be sufficient to consider the method as validated). One can also use a combination of approaches, e.g., retrospective assessment followed by prospective validation to close gaps in the dataset. In this paper, the term "validation" mostly refers to prospective validation unless otherwise specified.

3 Sources of candidate methods for validation

Method development projects encompass validation activities to various degrees and raise different issues:

- *Methods developed in research labs to answer a defined scientific question in a specific context:* Data generated by such methods can be used to add weight-of-evidence in a specific case but may not have high relevance otherwise. Thus, the need to demonstrate reproducibility is low if there is no need for prospective routine testing using the given method. An OECD project is on-going to develop guidance on the use of research data to increase confidence for regulatory risk assessment[2].
- *Publicly-funded research aimed at developing methods for regulatory applications, e.g., EURION cluster projects (€50 million – the largest funding of endocrine disrupting chemical research in Europe[3]), where validation operations or costs, as described in this paper, are not sufficiently covered:* As a consequence, methods developed under these research programs may, in most cases, be insufficiently validated at project completion; for example, reproducibility over time or a documented laboratory-to-laboratory transferability study may often be missing. However, there are a couple of examples of projects that have gone a little further into validation. In particular, the ERGO project aimed specifically at updating and optimizing existing OECD Test Guidelines with validated endpoints indicative of thyroid hormone system disruption (Holbech et al., 2020). During the course of the project, multiple methods were developed and optimized following validation principles in different laboratories, and the full validation study of two methods was initiated during the final months of this project. Also, some of the other EURION projects have been transferring methods to other laboratories and testing unblinded chemicals.

[1] Mutual Acceptance of Data: The 1981 OECD Council Decision on the Mutual Acceptance of Data in the Assessment of Chemicals (revised in 1997) states that test study data generated in any adhering country in accordance with OECD Test Guidelines and Principles of Good Laboratory Practice (GLP) shall be accepted in other adhering countries for assessment purposes and other uses relating to the protection of human health and the environment. (https://www.oecd.org/chemicalsafety/testing/mutualacceptanceofdatamad.htm).

[2] A guidance document that provides links to resources, good practices, and a case study example of using research data in regulatory risk assessments is currently under development at OECD and is expected to be finalized in 2025. A January 2024 webinar organised by the European Commission's Joint Research Centre, who lead the OECD project, includes an overview and presentations to give an idea of the challenges and opportunities for using research data in regulatory assessments (https://joint-research-centre.ec.europa.eu/events/online-webinar-good-practices-and-resources-improve-utility-research-data-regulatory-assessments-2024-01-31_en).

[3] https://eurion-cluster.eu

- *Privately-funded method development for commercialization, where validation is a real added value to a developer company's business model:* Such methods may contain protected elements; "check points" by independent parties may not be part of the process, but will be important to maintain trust.
- *Methods developed by chemical industry (e.g., pharmaceutical, cosmetic, or agrochemical companies) for in-house use in molecule discovery, screening for activity/efficacy for specific applications, and/or unwanted effects:* Such methods typically remain company assets and are not always validated for a broader implementation. The competitive context in new molecule development and the fact that these methods can be used in regulatory dossiers (e.g., by individual companies) without an extensive external validation is not an incentive for their broad validation. These methods are certainly an untapped source of new approach methods for use in a regulatory context.

4 What are the benefits of validation and where are the challenges?

We see the following main benefits associated with having validated methods:
- Prospective or retrospective validation builds confidence in the use of a method and consequently is a very important step towards regulatory acceptance.
- Validated methods have mature and well described procedures that help future study conductors to carry out the tests properly for regulatory purposes.
- Validated methods have a status (i.e., building block) that can be acknowledged by a broad community of potential users, even beyond the originally intended users, making it possible to extend the scope for other purposes.
- Validated methods allow to set regulatory data requirements, e.g., the systematic testing of new chemicals. It would not be possible to require the systematic testing of hundreds or thousands of chemicals with unvalidated methods, especially if this would have regulatory consequences like classification and labelling.
- Validation of a method is a prerequisite to adoption by the OECD as a Test Guideline, making the data generated (under GLP) in routine practice mutually accepted across OECD member and adhering countries, thus reducing duplicative testing.
- Results obtained with validated methods should also be seen as a way to increase reliability and trust toward scientific results in general.

Over the years, the following challenges have been identified:
- Significant resources are needed to carry out a prospective validation study to completion; these are often under-estimated from the outset. It is important to consider that the outcome of validation activities benefits many users (method developers, contract research organizations, industries, regulators), thus there should be ways to share responsibilities, including funding.
- Validation is typically non-appealing to the scientific community, which does not see the benefits of reproducing results multiple times, since the work does not usually result in publications in the peer-reviewed literature. Moreover, the lack of funding for validation studies makes it especially challenging for academic laboratories to participate in such efforts. This lack of interest and funding produces a disconnect between academic method developers and deployment of the discovery and associated technology.
- Typically, a fraction of method developers (those who also have a foot in investment and business) will seize the opportunity to validate new approach methods; this trend is emerging and, whilst it may provide one avenue for validating methods, it can also pose challenges for transparency in case of protected elements and/or possible situations of abuse of monopoly.
- Although implementation of validation offers some flexibility, there is a process to follow, which requires scientific and technical rigor and where shortcuts can be counter-productive.
- Quite often, there is not enough knowledge in the Validation Management Team on how to design and conduct a validation study, leading to various deficiencies that could have a negative impact on acceptance. Clear guidance on these processes is currently lacking.
- Given the innovation rate in biotechnologies and the global incentives to move away from animal testing, a virtuous cycle would require agility of chemical regulations for the rapid uptake of new validated methods, which would financially incentivize validation of innovative methods and thus facilitate the evolution of regulatory approaches to the safety evaluation of chemicals. However, although the regulatory need for new approaches for chemical safety testing is obvious, we observe reluctance to invest in validation of innovative methods in the absence of clear indications of how they would be used, in combination with other elements/methods, to provide a useful solution for regulatory decisions. As a consequence, the virtuous cycle is impeded.

5 What are the learnings from validation studies carried out in the last two decades?

Examples of efficient validation studies such as those performed for skin and eye irritation, and skin sensitization methods are characterized by:
1) Method developers proposing mature protocols,
2) Substantial lists of reference chemicals,
3) Clear data reporting and interpretation procedures,
4) Involvement of statisticians at the validation study design stage,
5) Support from a Validation Management Team (including experts in the method, in the technology used, and in validation),
6) A study design carefully prepared before starting the work, including an evaluation of the resources needed (operational, human, and financial).

One observation is that when there is urgent and clear regulatory and industry need to demonstrate safety, and the science/tech-

nique is mature, the process of method validation and acceptance as a whole can run efficiently and smoothly. However, it should be kept in mind that the above-mentioned examples relate mostly to situations where there was a broad list of well-characterized reference chemicals curated from *in vivo* data, and more particularly human data (i.e., known human irritants or sensitizers). In contrast, when validating methods for endocrine disrupters, as we discussed at the ED Forum (EU, 2022), the ability to identify a suitable range of reference chemicals can be a real limiting factor, since there are often fewer chemicals currently known to be active or known to be inactive in relation to specific endocrine modes of action.

Furthermore, the transfer of a method to a "naïve" laboratory is usually found to proceed more smoothly where there is a certain threshold of familiarity with the assay technology (e.g., proficiency in cell culture for a cell-based assay, device and equipment used in the assay); involving a laboratory that has limited proficiency in the techniques/equipment used poses a challenge for the time efficiency of the validation.

6 How to start on a good basis? The necessity to agree on what "readiness" means in the context of validation

Where method development ends and validation begins is not a bright line. A method that is a candidate for validation requires: i) a detailed method description presenting the biological or mechanistic basis underpinning the relevance of the method for the intended purpose, ii) a protocol or protocols to describe steps in the procedure (i.e., standard operating procedures (SOP)), iii) some data generated with the method with a few chemicals of known activity, where the expected result is clear (positive control, negative control), to show that the method responds as expected, and iv) some reference chemicals with a good dose-response relationship, to which repetitions of the method in different laboratories can be compared. In the case of cell-based assays, for example, the SOP should precisely state the cell type(s), growth conditions, and number of cell passages that can be used. The SOP may evolve along the process.

Data reporting templates in easily readable/usable format should be developed to be used by all the participating labs, facilitating comparison of reproducibility across experiments. They should be quality-controlled prior to their deployment in a validation study.

In an ideal situation, all these elements would be available upstream of the validation. However, this is rarely the case, and thus the acquisition of these elements extends the duration of the process. Method developers need to understand the expectations of method readiness before engaging time and resources in further validation including transferring the method to other laboratories. For the development and validation of *in vitro* methods the OECD guidance documents on validation and international acceptance (GD 34, OECD, 2005) and Good In vitro Method Practices (GIVIMP, OECD, 2018) are the reference documents to follow.

Even in the best-case scenario, some improvements or optimizations of the SOP are likely to be required before a mature, relatively stable SOP is defined. Method developers can self-evaluate to which extent their method is ready, for example, via a web tool such as readEDtest (Crouzet et al., 2023). This upstream in-house preparatory work by the method developer greatly facilitates downstream validation efforts, reducing the time taken to transfer the method to another laboratory.

7 After a method is considered "ready" for validation, a clear and optimized validation plan is necessary

In addition to the SOPs and reporting templates, before embarking on a validation study, a "validation plan" should be developed, including not only technical but also economic, logistical, legal, and ethical considerations. In particular, the availability of test method components (e.g., cells, tissue, organisms, equipment) and the identification of any protected elements under fair, reasonable, and non-discriminatory (FRAND) terms and conditions (OECD, 2019a) is essential to ensure the method can be used at a commercial scale and eventually be included in an OECD Test Guideline (if the validation is successful). For example, when an assay is initially developed using primary cells, broad commercial availability for a validation study and future routine testing may be a problem. Scale-up via immortalization of the cell line may be a pre-requisite to ensure that it will be accessible, and the assay will be deployable. It should be ensured that modified organisms such as transgenic fish can be made available to all participating laboratories (e.g., through material transfer agreements). Generally, the developer should make sure that test system components and reagents can be available commercially or from, e.g., certified biological resource centers.

The human and financial resources required in validation activities need to be itemized in sufficient detail to facilitate cost estimation ahead of the validation. A detailed and realistic budget is paramount to attract and mobilize dedicated funding.

In parallel, the method developer or validation study sponsor needs to establish a validation management team including expertise in the organization of validation studies, the assay technology, relevant biology/physiology, reference data/chemicals, and statistics.

8 Opportunities to reduce costs of validation

The greatest costs of a validation study are usually the human resources in terms of the laboratory technicians and other experts involved, which are difficult to reduce. The PEPPER platform[4], a public-private partnership for the validation of endocrine disruption-related methods, has some years of experience with several validation studies currently on-going in parallel. The annual budget of the platform has been broken down into 3 categories

[4] https://ed-pepper.eu/en/

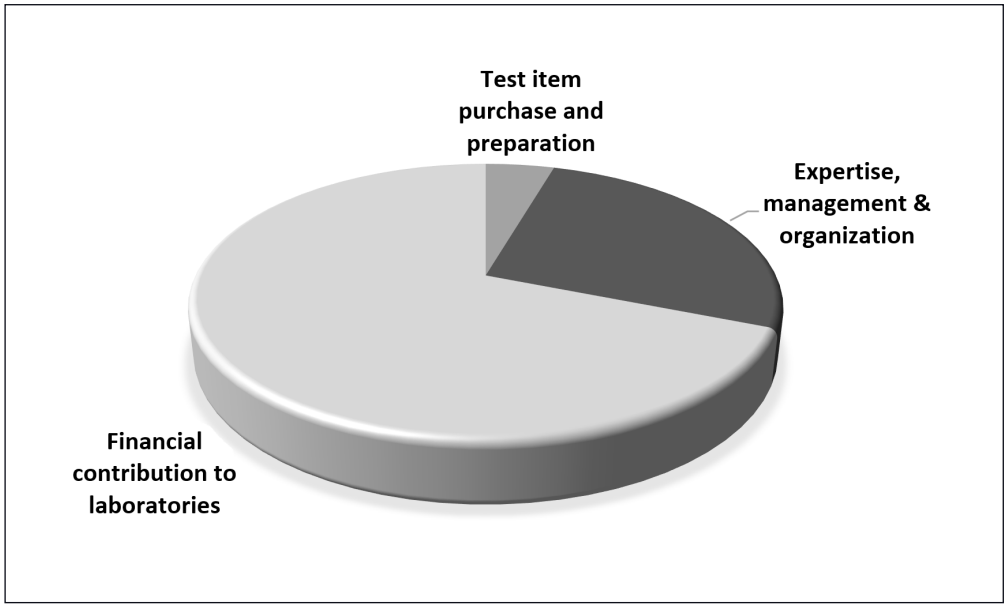

Fig. 1: Representation of the distribution of validation costs, as experienced by PEPPER

(see Fig. 1). The main one is the financial contribution to laboratories, even though laboratories are not 100% supported by PEPPER. The second is the expertise and organizational costs, which is mostly internal staff cost of PEPPER, as well as external assistance (e.g., for statistics and chemical (test item) selection) but not including the drafting of validation study reports or the peer review process. The third category is on purchase and shipping of these test items (after blind coding).

Efficiencies can be gained in a number of areas, and a few suggestions for consideration are given below.

- *Reagents, test chemicals:* A central repository of reference chemicals for different hazard endpoints in each global region/continent where validation activities are on-going to limit issues with customs, and the distribution of samples of stock solutions in amounts needed; blinding can be done by the repository as appropriate. Such a repository should be sponsored by multiple partners (public, private) and seen as a common good. It should be a long-term supported endeavor.
- *Analytical chemistry:* The need for analytical verification of certain aspects, once agreed, could be centralized to one laboratory if the other laboratories cannot do it.
- *Training:* The development of online training should be exploited as much as possible, as it is re-usable, can be augmented by the method developer, and can benefit multiple users at limited costs;
- *Data management:* Streamlining the data flow and data analysis with optimized reporting templates and a well-designed integrated flow between laboratories and sponsors with automated controls would save a lot of time in quality control and data handling.

Depending on the assays considered, the above suggestions may have more or less impact on the overall costs of the validation study.

9 Giving more weight to well conducted and documented transferability studies

The practice of ring-trials typically involving the testing of a substantial number of chemicals across more than three laboratories is a very resource-intensive and costly activity. It not only demonstrates that a method is reproducible but also brings a demonstration of laboratory proficiency across a large community of potential users. It builds a high degree of confidence in the strengths and limitations of a method. However, it can be questioned whether the efforts and logistics involved are always proportionate to the value added to the method.

When considering efficiencies in demonstrating between-laboratory reproducibility, the need for extensive ring trials with more than three labs is currently discussed (OECD, 2023), and particularly, to what extent transferability studies to just one or two labs that are well planned, monitored and reported, could be sufficient to build confidence and trust that a method can be reproduced with high fidelity by another laboratory without bias.

Successful and well-documented transferability studies may serve third parties wanting to adopt a new method or technology. This reduces the cost of validation by reducing the amount of testing conducted overall but also by not having to dedicate significant resources to the coordination/training of many laboratories, which may have different proficiency levels at the outset. Training material can be developed during a transferability study and reused by third party laboratories for proficiency acquisition. Having the method successfully transferred to one or two laboratories may be enough to demonstrate reproducibility outside the developer's laboratory. Based on the results of these transferability studies, a list of proficiency chemicals should be established to help future labs implementing the method to self-assess their proficiency level.

The issue of ring-trial practice and cost-benefit analysis was discussed at the ED forum and is still a matter under discussion in the broader regulatory science community. While some are of the opinion that successful ring-trials with multiple laboratories bring confidence that the method is not only reproducible but also immediately deployable across many facilities, others are of the opinion that validation is an incremental effort where the key milestone is the successful transfer of the method, including the blind-testing of chemicals. In reality, everyone is correct, but the latter tend to dissociate broad proficiency acquisition from pure method validation while the former think that there is an increased risk of failure discovery during implementation if no ring-trial has been carried out.

10 What will new approach methods look like and how will processes ensuring relevance and reliability need to adapt?

Many different solutions to chemical safety assessment aiming to reduce reliance on testing in animals are in development. Current attempts to predict systemic toxicity for a range of complex hazard endpoints are focusing on understanding molecular and cellular changes leading to toxicity at the tissue or organ level using combinations of methods in an integrated fashion. This is particularly relevant to endocrine disrupter identification where the focus is on understanding the mode of action of a chemical that leads to an adverse effect. In the future, information sources are expected to be used together to generate an output, with data being combined using a predictive model/algorithm that results in meaningful information that can be used by regulators. In this context, it is reasonable to reflect and rethink how much value and relevance is given to the predictive capacity of individual methods that are to be used as building blocks for a predictive model (defined approaches[5]).

In parallel, complex methods (e.g., microphysiological systems, organoids) that more closely recapitulate endocrine physiology are being developed and will require time for technology transfer and uptake. The increased sophistication of new approach methods and often the scarcity of publicly available reference data may limit the number of key players with access to costly equipment, proprietary data, and necessary skills to propose and implement innovative solutions. Having fewer key players at the deployment stage (around and after validation) should not be a barrier to the acceptance of the technology. Although the situation represents a difficulty for the recruitment of suitable laboratories for method transfer, proficiency acquisition in the larger community of potential users is an important issue but can be gradual as the new methods deploy and should generally be kept a separate issue from the validation and performance evaluation of a method. From experience of validation studies coordinated by EURL ECVAM, no method failed the between-laboratory reproducibility once the within-laboratory reproducibility was considered sufficient.

11 Validation, a shared responsibility

This was discussed at the ED Forum. The authors consider that validation is a common good that brings benefit to a broad community and thus it should be a shared responsibility, especially the funding part, which is sorely lacking. There were diverse opinions on the financial responsibilities among panel members at the ED Forum but broad agreement that both private and public funding sources should be involved.

Following that event, a number of other initiatives have been taken, in particular at the OECD, to raise awareness on the operational and financial aspects of validation among the stakeholder community (OECD, 2024, in prep).

Validation is not considered a part of method development. However, validation can lead to method improvement in laboratories and also comprehensive, standardized, and effective data reporting. Although capacity building of naïve laboratories can inform the future capacity to deploy the method for routine testing, it remains extremely time-consuming and separate from the intrinsic performance evaluation of a method.

Although validation may be perceived as a discrete task whose responsibility/involvement remains in the hands of validation bodies that are funded to perform that job, in our opinion responsibilities for validation should be shared by all concerned parties that benefit from it (method developers, contract research organizations, industry, animal welfare non-governmental organizations (NGOs), and regulators).

Industry, who primarily benefit from validated methods to fulfil their regulatory obligations, should invest in validation, especially in the case of NAMs, which are intended to reduce costs and the use of animals for them. Funding from industry, managed by independent entities, should be used for validation and peer-review purposes. This would be a win-win for all parties involved.

In the case of NAM validation, we discussed at the ED Forum in September 2022 that also animal welfare NGOs could contribute more, as their goal to reduce animal testing is clearly supported by validation of non-animal methods. Such funding already (partially) exists, for example from the PETA International Science Consortium, which supports the development of NAMs for regulatory testing.

In the end, protection of environmental and human health is of high public interest and, thus, it is reasonable to expect that public funding (e.g., from the EU) should also be allocated to validation. For example, this was clearly stated and required in EU grant agreements of the research projects in the EURION cluster. This will provide some financial support for validation at academic laboratories, but finally, the researchers working on these activities

[5] A defined approach consists of a fixed data interpretation procedure (DIP) (e.g., statistical, mathematical models) applied to data (e.g., *in silico* predictions, *in chemico*, *in vitro* data) generated with a defined set of information sources to derive a prediction. In contrast to the assessment process within integrated approaches to testing and assessment (IATA), which necessarily involves some degree of expert judgment, predictions generated with defined approaches are rule-based and can either be used on their own if they are deemed fit-for-purpose or considered together with other sources of information in the context of IATA (OECD, 2016).

also need to be rewarded in their own "currency" – publications. Peer-reviewed journals should ensure that the topic of method development and validation is within the scope of articles accepted for publication, which is crucial for the careers of academic researchers.

12 Conclusion

As identified during the ED Forum, there is an increased awareness of the need to ensure test method validation, but the engagement of stakeholders in taking more responsibilities and funding calls for validation still needs to be improved. As validation becomes a shared responsibility, we now see various stakeholders organizing themselves to coordinate validation activities, including private entities, public-private partnerships, and publicly funded research programs. There is more acceptance that everyone should play a role and participate in funding. A recent OECD survey related to the cost of validation has come up with a range of cost estimates from 200 to 500 kEUR (not necessarily including manpower), depending on the complexity of the assay. However, it is not always clear to what extent this represents the true costs since figures may be grossly underestimated where labor costs are not fully incorporated and "hidden costs" such as the preparation of reports and expert analysis are not included in the budget. Based on these costs, it is a positive evolution that more stakeholders share the costs of validation activities and that the level of expertise required in managing validation activities is also very well acknowledged. We acknowledged in our discussions the issue of ring-trials with their costs and their benefits and the need to find the right solution (costs/resources involved and benefits obtained) for each situation. Indeed, these costs have to be balanced against the benefits of ultimately having validated and accepted methods, allowing Mutual Acceptance of Data. The costs are also relatively minor when compared to the cost of "no validation". OECD estimated the savings that come from the Mutual Acceptance of Data to be at the level of 309 million EUR each year (OECD, 2019b). It is therefore clear that having validated methods is a profitable investment, not just a cost.

References

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature 533*, 452-454. doi:10.1038/533452a

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

Crouzet, T., Grignard, E., Brion, F. et al. (2023). ReadEDTest: A tool to assess the readiness of in vitro test methods under development for identifying endocrine disrupters. *Environ Int 174*, 107910. doi:10.1016/j.envint.2023.107910

EU (2022). Report of the 4th Annual Forum on Endocrine Disrupters. Exchanging knowledge, identifying challenges, building synergies, 21-22 September 2022. Luxembourg: Publications Office of the European Union. doi:10.2779/71861

Holbech, H., Matthiessen, P., Hansen, M. et al. (2020) ERGO: Breaking down the wall between human health and environmental testing of endocrine disrupters. *Int J Mol Sci 21*, 2954. doi:10.3390/ijms21082954

OECD (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. *OECD Series on Testing and Assessment, No. 34*. OECD, Paris. https://one.oecd.org/document/env/jm/mono(2005)14/en/pdf

OECD (2016). Guidance Document on the reporting of defined approaches to be used in integrated approaches to testing and assessment. *OECD Series on Testing and Assessment, No. 255*. OECD, Paris. https://one.oecd.org/document/env/jm/mono(2016)28/en/pdf

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

OECD (2019a). Guiding Principles on Good Practices for the Availability/Distribution of Protected Elements in OECD Test Guidelines. *OECD Series on Testing and Assessment, No. 298*. OECD, Paris. https://one.oecd.org/document/env/jm/mono(2019)14/en/pdf

OECD (2019b). Saving Costs in Chemicals Management: How the OECD Ensures Benefits to Society https://www.oecd.org/chemicalsafety/saving-costs-in-chemicals-management-9789264311718-en.htm

OECD (2023). Report on the WNT Workshop how to prepare the Test Guidelines Programme for emerging technologies. *OECD Series on Testing and Assessment, No. 378*. OECD, Paris. https://one.oecd.org/document/ENV/CBC/MONO(2023)14/en/pdf

Piersma, A. H., Burgdorf, T., Louekari, K. et al. (2018). 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

van der Zalm, A. J., Barroso, J., Browne, P. et al. (2022). A framework for establishing scientific confidence in new approach methodologies. *Arch Toxicol 96*, 2865-2879. doi:10.1007/s00204-022-03365-4

Disclaimer

The views expressed in this article are those of the authors and do not necessarily represent the position of their affiliated institution.

Conflict of interest

EG, PH and SM are employees of a structure that supports the validation of methods. The other authors declare no conflict of interest.

Data availability

This paper does not contain any scientific data to be shared.