



Advances in the Development and Validation of Test Methods in the United States

Warren M. Casey

*Director, NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)
Executive Director, Interagency Coordinating Committee on the Validation of Alternative Toxicological Methods (ICCVAM)
Division of the National Toxicology Program, National Institute of Environmental Health Sciences, RTP, NC, USA*

(Received October 23, 2015; Revised December 30, 2015; Accepted January 7, 2016)

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) provides validation support for US Federal agencies and the US Tox21 interagency consortium, an interagency collaboration that is using high throughput screening (HTS) and other advanced approaches to better understand and predict chemical hazards to humans and the environment. The use of HTS data from assays relevant to the estrogen receptor signaling data pathway is used as an example of how HTS data can be combined with computational modeling to meet the needs of US agencies. As brief summary of US efforts in the areas of biologics testing, acute toxicity, and skin sensitization will also be provided.

Key words: Alternative, Tox21, HTS, US, Toxicity, Validation

INTRODUCTION

Considerable worldwide interest exists in developing an understanding of how biologic pathways become engaged in toxicity responses. The intent is that knowledge about these pathways would lead to the development of predictive, integrated testing strategies that likely would combine in silico approaches and multiple in vitro or high-throughput assays with limited, targeted testing in laboratory animals. There are various statistical approaches to assembling components of a 'most predictive' test battery as well as different statistical models for integrating all relevant information and assay results. Both the batteries and models as alternative methods may warrant consideration by U.S. agencies. In addition, numerous efforts are underway to develop in silico and in vitro approaches for screening and

prioritization of chemicals for testing. Ultimately, the application of these alternative methods may differ (regulatory testing for safety evaluation or risk assessment compared with screening or prioritization), and as such, the concept of validation in the context of use needs to be considered. US agencies intend to engage the public in discussions about the concept of validation and its role in this new paradigm. In the future, more complex batteries of tests that evaluate systemic toxicity will need to be assembled. As a consequence, US agencies will place less emphasis on replacement of in vivo toxicity endpoints with a single alternative assay and more emphasis on assembling batteries of assays and including in silico approaches to interpret and analyze data from such batteries. In addition to changing the concept of validation, agencies anticipate that reevaluating standing policy in response to new data could substantially affect the number of animals used for regulatory safety testing.

Correspondence to: Warren M. Casey, Director, NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), Executive Director, Interagency Coordinating Committee on the Validation of Alternative Toxicological Methods (ICCVAM), Division of the National Toxicology Program, National Institute of Environmental Health Sciences, RTP, NC, USA
E-mail: warren.casey@nih.org

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ICCVAM. In 2000, the ICCVAM Authorization Act established the Interagency Coordinating Committee on the Validation of Alternative Toxicological Methods (ICCVAM) as a permanent interagency committee of the NIEHS under NICEATM (1). ICCVAM is composed of representatives from 15 U.S. Federal agencies (Consumer Product Safety Commission, Department of Agriculture, Department of the Interior, Department of Transportation, Environmental Protection Agency, Food and Drug Administration, and Occupational Safety and Health Administration, Agency for

Toxic Substances and Disease Registry, National Institute for Occupational Safety and Health, National Cancer Institute, National Institute of Environmental Health Sciences, National Library of Medicine, National Institutes of Health, Department of Defense, and Department of Energy) that require, use, generate or disseminate toxicological and safety testing information. ICCVAM's mission is to "facilitate development, validation, and regulatory acceptance of new and revised regulatory test methods that reduce, refine, and replace the use of animals in testing while maintaining and promoting scientific quality and the protection of human health, animal health, and the environment." ICCVAM relies on stakeholders to carry out research, development, and validation, providing guidance to test method developers where needed. As such, ICCVAM does not carry out validation studies on behalf of test method developers, but evaluates alternative toxicological test methods submitted by stakeholders and reviews validation data, organizes expert peer reviews of promising methods, and makes recommendations on the use of reviewed test methods to appropriate U.S. Federal agencies. To optimize utilization of resources and avoid duplication of effort, ICCVAM coordinates test method evaluation activities with ICATM partner organizations (currently Korea, Japan, European Union, and Canada). International collaboration is also facilitated through ICCVAM agency participation in the Organization for Economic Cooperation and Development (OECD). Broader engagement with the scientific community and stakeholders is achieved through focused workshops, webinars and face-to-face forums.

NICEATM. The NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is part of the National Institute of Environmental Health Sciences (NIEHS). NICEATM activities support the development and evaluation of new, revised, and alternative methods to identify potential hazards to human health and the environment, with a focus on replacing, reducing, or refining animal use. NICEATM activities include:

- Conducting and publishing analyses and evaluations of data from new, revised, and alternative testing approaches
- Providing information to test method developers, regulators, and regulated industry through this Website and other communications and by organizing workshops and symposia on topics of interest
- Supporting the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM)

It is important to note the distinction in the roles of ICCVAM and NICEATM; ICCVAM is responsible for determining the relevance and priority of test method nominations, whereas NICEATM conducts their scientific review by making use of its competencies in validation study design, computational toxicology, chemoinformatics, data management and data analysis.

Tox21. Tox21 is a multiagency collaboration that includes NIEHS-NTP, the NIH's National Center for Advancing Translational Sciences, the EPA's National Center for Computational Toxicology, and the Food and Drug Administration. The objective of this partnership is to shift the assessment of chemical hazards away from traditional laboratory animal toxicology studies to an evaluation based on target-specific, mechanism-based, biologic observations largely obtained by using in vitro assays, with the ultimate aim of improving risk assessment for humans and the environment. Tox21's strategy is to use in vitro high-throughput and high-content screening technologies to test a broad variety of approximately 10,000 substances and consider data from those screens collectively to assess effects on biologic pathways related to toxicity. Data from Tox21 testing will be used to develop a better understanding of these toxicity or 'adverse outcome' pathways (AOPs), ideally enabling the eventual use of in vitro assay data to predict the adverse effects of chemical exposures in vivo. More information about Tox21 can be found on the NTP website (<http://ntp.niehs.nih.gov/go/tox21>).

Beginning in 2013, NICEATM expanded its activities beyond ICCVAM. In her Environmental Health Perspectives editorial NIEHS and NTP Director Linda Birnbaum enlarged the scope of NICEATM to include providing bioinformatic and computational toxicology support to the interagency Tox21 effort (2). The intent was to position NICEATM in a bridging role between ICCVAM and Tox21. These organizations are working together to promote more efficient approaches to develop and validate methods that employ the 3Rs, and they are exploring options that would allow these methods to be expeditiously accepted by regulators and used by stakeholders.

ICATM. The International Cooperation on Alternative Test Methods (ICATM) ICATM was established in 2009 to promote consistent and enhanced voluntary international cooperation, collaboration, and communication among national validation organizations. The goals of ICATM are to: 1) ensure the optimal design and conduct of validation studies; 2) ensure high-quality independent scientific peer reviews of alternative test methods; 3) ensure consistent and transparent stakeholder involvement; 4) achieve greater efficiency and effectiveness by internationally leveraging limited resources and avoiding duplication of effort; and 5) support the timely international adoption of alternative test methods. This cooperation enables scientifically valid alternative methods or strategies to be more readily accepted worldwide for regulatory use. In addition to ICCVAM, the current ICATM partner organizations are The European Union Reference Laboratory for Alternatives to Animal Testing (EURL_ECVAM); The Japanese Center for the Validation of Alternative Methods (JaCVAM); Health Canada's Environmental Health Science and Research Bureau;

and The Korean Center for the Validation of Alternative Methods (KoCVAM).

SCREENING AND PRIORITIZATION OF ENDOCRINE DISRUPTORS

The US EPA's Endocrine Disruptor Screening Program (EDSP) was established in 1999 for the purpose of evaluating potential risk of endocrine disruption in humans and wildlife from exposure to pesticide chemicals and drinking water contaminants. To screen chemicals for potential endocrine bioactivity, EPA developed a battery of five in vitro and six in vivo Tier 1 screening assays¹. In response to the US National Academy of Sciences report, Toxicity Testing in the 21st Century (3), EPA began a transition from existing EDSP test methods towards utilizing more rapid, cost-effective computational models and high-throughput assays. EPA's Toxicity Forecaster (ToxCast) program generates data on ~2000 chemicals of interest to the EPA. ToxCast uses high-throughput screening methods and computational toxicology approaches to rank and prioritize chemicals. EPA's Endocrine Disruption Screening Program (EDSP) is now using ToxCast data to rank and prioritize chemicals based on their potential ER bioactivity, serving as an alternative for some of the existing EDSP Tier 1 in vitro and in vivo assays (4).

In order to use new computational toxicology approaches in the existing EDSP screening and testing framework, the new approaches must be validated and perform as well or

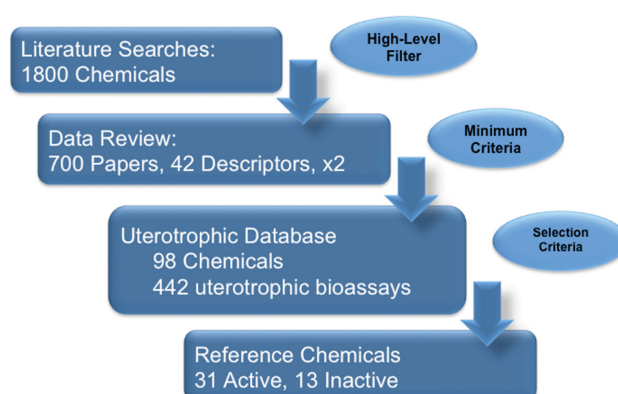


Fig. 2. Flow diagram illustrating the curation of the uterotrophic database (UTDB) and identification of high-quality guideline like (GL) studies.

better than existing methods currently in practice. The validation of the new screening approach for ER bioactivity is briefly described here. High-throughput in vitro screening assay data from 16 different ER-based assays in the EPA's ToxCast program were integrated into a computational model of the ER pathway bioactivity (5), and model performance was compared with in vitro and in vivo reference chemicals identified from the peer-reviewed literature (6). Details of the in vitro assays have been previously published (7,8) and are also described on EPA's ToxCast website (<http://actor.epa.gov/dashboard/>). The combination of biochemical and cell-based in vitro assays relies on different technolo-

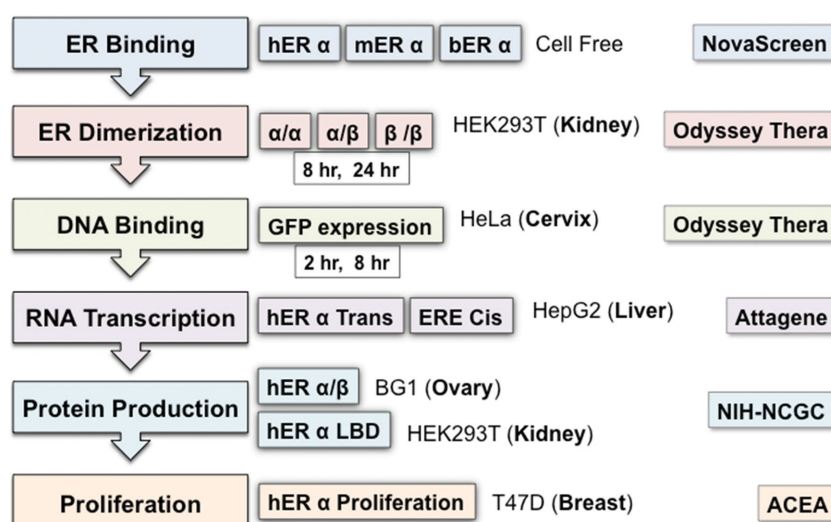


Fig. 1. 16 ER pathway assays from ToxCast. These include three cell-free biochemical radioligand ER binding assays from NovaScreen (human, mouse, bovine); a set of three protein complementation assays from Odyssey Thera that measure formation of ER dimers and test for activity against ER-alpha and ER-beta, (each homo- and hetero-dimer measured at two separate times for a total of six assay readouts); two Odyssey Thera assays measuring interaction of the transcription factor with DNA; two transactivation assays from Attagene measuring RNA transcript levels; two transactivation assays measuring reporter protein level readouts, conducted at NIH-NCGC; and an ER-sensitive cell proliferation assay from ACEA. The assay sources refer to the company or laboratory where the assays were performed. Cell line and tissue of origin are also provided for each assay.

gies and probes different key events in the ER signaling pathway (Fig. 1). Every in vitro assay is potentially subject to technology-specific interference that can be mistakenly interpreted as bioactivity. However, combining data from multiple orthogonal assays and integrating data in a network model of the entire ER pathway allows for detection of false positives and a more confident assessment of the “true” in vitro estrogenic bioactivity of the tested chemical. The ER model scores generated using in vitro data were compared to results obtained from high quality in vivo stud-

ies. In vivo reference chemicals were established from a literature search of short-term rodent assays that were methodologically similar to the OECD and EDSP Tier 1 battery uterotrophic tests (6). Chemical data from curated uterotrophic studies were considered with two levels of stringency (Fig. 2). First, a chemical was considered positive for potential in vivo estrogen agonist bioactivity if a significant increase in uterine weight among treated animals was reported and negative if no significant increase in uterine weight was reported in any guideline-like study. Second,

Table 1. *In vivo* estrogen agonist reference chemicals with at least two independent active or inactive guideline-like uterotrophic studies identified from a comprehensive literature review and curation (reproduced from Kleinstreuer *et al.* 2015 with permission)

CASRN	Name	Bioactivity
57-91-0	17alpha-Estradiol	Active
57-63-6	Ethinyl Estradiol	Active
56-53-1	Diethylstilbestrol (DES)	Active
50-28-2	Estradiol	Active
474-86-2	Equilin	Active
53-16-7	Estrone	Active
50-27-1	Estriol	Active
72-33-3	Mestranol	Active
17924-92-4	Zearalenone	Active
1478-61-1	Bisphenol AF	Active
446-72-0	Genistein	Active
68-22-4	Norethindrone	Active
58-18-4	Methyltestosterone	Active
77-40-7	Bisphenol B	Active
80-05-7	Bisphenol A	Active
104-43-8	4-Dodecylphenol	Active
521-18-6	Dihydrotestosterone	Active
131-55-5	Benzophenone-2	Active
140-66-9	4-(1,1,3,3-Tetramethylbutyl)phenol	Active
789-02-6	o,p'-DDT	Active
599-64-4	p-Cumylphenol	Active
5153-25-3	Benzoic acid, 4-hydroxy-, 2-ethylhexyl ester	Active
80-46-6	4-(1,1-Dimethylpropyl)phenol	Active
131-56-6	2,4-Dihydroxybenzophenone	Active
80-09-1	Bisphenol S	Active
72-43-5	Methoxychlor	Active
94-26-8	Butylparaben	Active
98-54-4	p-tert-Butylphenol	Active
556-67-2	Octamethylcyclotetrasiloxane	Active
520-18-3	Kaempferol	Inactive
84-74-2	Dibutyl phthalate	Inactive
84-61-7	Dicyclohexyl phthalate	Inactive
84-75-3	Dihexyl phthalate	Inactive
51630-58-1	Fenvalerate	Inactive
103-23-1	Bis(2-ethylhexyl)hexanedioate	Inactive
117-81-7	Bis(2-ethylhexyl)phthalate	Inactive
1461-22-9	Tributylchlorostannane	Inactive
1912-24-9	Atrazine	Inactive
61-82-5	Amitrole	Inactive
84-66-2	Diethyl phthalate	Inactive
87-86-5	Pentachlorophenol	Inactive
99-96-7	4-Hydroxybenzoic acid	Inactive

only chemicals tested in two or more guideline-like uterotrophic studies were used. Chemicals that resulted in a significant increase in uterine weight in two or more independent guideline-like studies were considered positive, while those chemicals that showed negative results in all studies were considered negative. This second subset of chemicals were referred to as in vivo reference chemicals since study results were reproducible (Table 1). For the 43 in vivo reference chemicals, the overall balanced accuracy of the ToxCast ER model was 86%, providing a high degree of confidence in ToxCast ER model predictions and demonstrating the utility of using these computational tools to meet the intended objectives which were to: 1) to contribute to the weight of evidence evaluation of a chemical's ER agonist activity; and 2) to provide an alternative source of data for specific Tier 1 endpoints measuring in vitro and in vivo ER interaction.

OTHER ICCVAM ACTIVITIES

ICCVAM has identified several projects where the science has advanced substantially and where there is a reasonable likelihood of implementation into regulatory use within the next 5 years.

Regarding biologics in general and *Leptospira* vaccine specifically, the US Department of Agriculture's Animal and Plant Health Inspection Service is committed to decreasing the number of hamsters used in *Leptospira* vaccine potency testing. In addition, USDA is developing a mechanism to use fewer hamsters in the maintenance of the *Leptospira* challenge cultures. The effects of this change on hamster usage will be monitored over the next 5 y. A 25% to 30% decrease in animal usage by the industry is anticipated.

In the realm of acute oral and dermal toxicity testing, the US Environmental Protection Agency is evaluating the relative contribution of acute and dermal LD50 tests in providing information related to hazard labeling and the use of personal protective equipment. In addition, ICCVAM is investigating the utility of in vitro assays (for example, 3T3 NRU) for predicting oral LD50 values.

Another priority area of ICCVAM for the next 5 years is skin sensitization. Great strides have been made in putting together a battery of in silico, in chemico, and in vitro tests for the assessment of skin sensitization. Such nonanimal alternative assays have been developed for the various key events in the adverse outcome pathway of skin sensitization. With support from the Environmental Protection Agency, Consumer Product Safety Commission, and Food and Drug Administration, ICCVAM has announced the development of a plan for the evaluation of alternative skin sensitization test methods and testing strategies. Activities proposed as part of the plan include collaboration with international partners to support ongoing development and validation of in vitro skin sensitization test methods; evaluation of alter-

native test method and testing strategy submissions for skin sensitization; and promotion of validated methods through workshops, webinars, and guidance documents. The newly established ICCVAM Skin Sensitization Working Group is reviewing and evaluating the public's response to the plan and will advance recommendations for appropriate ICCVAM activities for the next several years. More information about the ICCVAM plan to evaluate alternative skin sensitization test methods can be obtained on the NTP website (<http://ntp.niehs.nih.gov/go/40498>).

In the future, many ICCVAM efforts will be focused on In Vitro to In Vivo Extrapolation (IVIVE) Historically, toxicological testing has relied largely on the use of laboratory animals, and is therefore low throughput, high cost, and presents difficulties inherent to inter-species extrapolation. In comparison, high throughput screening (HTS) assays can accelerate and reduce the overall cost of the screening process and are being actively pursued through programs such as the Federal Tox21 consortium and the U.S. EPA's ToxCast project. However, before in vitro HTS assays can be used for risk assessment and regulatory decision-making, the ability of in vitro activity to predict in vivo dose-response relationships needs to be evaluated. There are many factors affecting in vitro to in vivo extrapolation. The primary question is how closely a bioactive chemical concentration from an in vitro assay corresponds to blood and tissue levels that cause adverse effects in vivo. ICCVAM and US Federal agencies will increasingly focus on developing and evaluating methods that enable the accurate characterization of concentration activity relationships in both in vitro and in vivo systems.

ACKNOWLEDGEMENTS

The work described here is a summary of many efforts across the US Federal Government. The contributions of David Dix, Patience Browne, Richard Judson, Nicole Kleinstreuer, and Rusty Thomas are acknowledged in particular.

REFERENCES

1. ICCVAM. (2000) Authorization Act of 2000. ICCVAM, **42**, 106-545.
2. Birnbaum, L.S. (2013) 15 years out: reinventing ICCVAM. *Environ. Health Perspect.*, doi:10.1289/ehp.1206292.
3. National Academy of Sciences. (2007) Board on environmental studies and toxicology. Toxicity testing in the 21st century: a vision and a strategy. The National Academies Press, Washington DC.
4. Browne, P., Judson, R.S., Casey, W.M., Kleinstreuer, N.C. and Thomas, R. (2015) Screening chemicals for estrogen receptor bioactivity using a computational model. *Environ. Sci. Technol.*, **49**, 8804-8814.
5. Judson, R.S., Magpantay, F.M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., Xia, M., Huang, R., Rotroff, D.M., Filer,

- D.L., Houck, K.A., Martin, M.T., Sipes, N., Richard, A.M., Mansouri, K., Setzer, R.W., Knudsen, T.B., Croffon, K.M. and Thomas, R.S. (2015) Integrated model of chemical perturbations of a biological pathway using 18 in vitro high throughput screening assays for the estrogen receptor. *Toxicol. Sci.*, **148**, 137-154.
6. Kleinstreuer, N.C., Ceger, P.C., Allen, D.G., Strickland, J., Chang, X., Hamm, J.T. and Casey, W.M. (2015) A curated database of rodent uterotrophic bioactivity. *Environ. Health Perspect.*, doi:10.1289/ehp.1510183.
7. Dix, D.J., Houck, K.A., Martin, M.T., Richard, A.M., Setzer, R.W. and Kavlock, R.J. (2007) The toxcast program for prioritizing toxicity testing of environmental chemicals. *Toxicol. Sci.*, **95**, 5-12.
8. Rotroff, D.M., Dix, D.J., Houck, K.A., Knudsen, T.B., Martin, M.T., McLaurin, K.W., Reif, D.M., Crofton, K.M., Singh, A.V., Xia, M., Huang, R. and Judson, R.S. (2013) Using in vitro high throughput screening assays to identify potential endocrine-disrupting chemicals. *Environ. Health Perspect.*, **121**, 7-14.