Workshop 5.13
Strategies for prioritising and streamlining the validation process

Lecture
Validation via weight-of-evidence approaches

Michael Balls and Robert Combes
FRAME, Nottingham, UK

It is not always possible, necessary, or even desirable, to evaluate the relevance and reliability of an *in vitro* or *in silico* test or testing strategy by comparing the predictions it provides with those obtained from an *in vivo* test or testing strategy. A weight-of-evidence approach aims to use all the available information that meets certain criteria, in a structured, systematic, independent and transparent review of relevance and reliability in relation to purpose, the outcome of which will be published in the peer-review literature. Crucial aspects include: The selection of the reviewers and the application of procedures to ensure independence and lack of bias; criteria for the selection of data; procedures for the collection of data and for data quality control; procedures for the differential weighing of various types of evidence, alone and in combination; criteria for test/strategy performance in terms of reliability and relevance in relation to purpose; and agreement on how the outcome should be expressed, published and otherwise made available. Examples will be given to illustrate where the weight-of-evidence approach has been used in the past and where it will need to be applied in the future.

Poster

ECVAM Key Area of Strategic Developments: Summary of ongoing activities

Sandra Coecke, Gerard Bowe, Juan Casado, Silvia Casati, Raffaella Corvi, Jan De Lange, Massimo Farina, Fernando Fernandez, Salvador Fortaner, Agnieszka Kinsner, Jens Linge, Siegfried Morath, Barbara Munaro, Nicholaos Parissis, Jessica Ponti, Pilar Prieto, Anna Price, Enrico Sabbioni and Erwin Van Vliet

ECVAM, Institute for Health and Consumer Protection, Joint Research Center, Ispra, Italy

Large numbers of animals are required for the regulatory testing of chemicals, cosmetics, biologicals, biomaterials, pharmaceuticals and other products. Alternative methods and new testing strategies, which could reduce, refine and replace the use of animals, are being developed and need to be validated and evaluated at European level.

The Key Area Strategic Developments relates to enabling technologies and activities (e.g. "omics" technology, high-throughput screening, GLP and GCCP, nanoparticle toxicology) and their use in development and validation of alternative methods for tackling new areas in toxicology and speeding up the validation process by introducing new technologies.

Omics technologies, fingerprints, pattern-based assessments and biomarker approaches are explored for use in regulatory toxicology in close co-operation with other international initiatives supported by in-house laboratory activities.

ECVAM is evaluating how high-throughput screening of substances with *in vitro* cellular systems could assist the validation process by running its own automated facility. This work is also serving the A-Cute-Tox FP6 Integrated Project.

Nanotechnologies and nanopartical toxicology are areas which are getting more on the political agenda's. ECVAM is involved in the FP6 STREP "ToxDrop" focussed on high content analyses of cell cultures. Furthermore, ECVAM's laboratory group is contributing to assessing the toxicity of organic and inorganic nanoparticles using *in vitro* cell and tissue cultures.

ECVAM continues to steer and contribute actively to advisory and guidance documents on GLP and GCCP crucial for the validation process.

By investing into enabling technologies we hope to reshape hazard identification in European legislation.

Lecture

A modular approach to the ECVAM principles on test validity

Thomas Hartung, Susanne Bremer, Silvia Casati, Sandra Coecke, Raffaella Corvi, Salvador Fortaner, Laura Gribaldo, Marlies Halder, Sebastian Hoffmann, Annett Janusch Roi, Pilar Prieto, Enrico Sabbioni, Laurie Scott, Andrew Worth and Valérie Zuang

ECVAM, Institute for Health and Consumer Protection, Joint Research Center, Ispra, Italy

The European Centre for the Validation of Alternative Methods (ECVAM) proposes to make the validation process more flexible while maintaining its high standards. The various aspects of validation are broken down into independent modules and the information necessary to complete each module is defined. The data required to assess test validity in an independent peer-review, not the process, is thus emphasised. Once the information to fulfil all modules is complete, the test can enter the peer-review process. In this way the between-laboratory variability and predictive capacity of a test can be assessed independently. Thinking in terms of validity principles will broaden the applicability of the validation process to a variety of tests and procedures, including the new generation of tests, new technologies (e.g. genomics, proteomics), computer-based models, and expert systems (e.g. (Q)SARs). Furthermore, this proposal aims to take into account existing information, defining this as retrospective validation, in contrast to a prospective validation study, which has been the predominant approach to date. This will allow the assessment of the test validity by completing the missing information via the relevant validation procedure: prospective, retrospective, catch-up validation, or combination of them.

Lecture

Optimising validation study designs by separating between-laboratory reproducibility and predictive capacity: The example of Skin Corrosion

Sebastian Hoffmann and Thomas Hartung
European Commission, JRC, ECVAM, Ispra, Italy

The modular approach to test validity assessment (Hartung et al. (2004), *Altern Lab Anim, 32*, 467-472) claims to be more flexible and efficient by defining validation by seven independent modules. In particular and in contrast to former approaches, the aspects of between-laboratory reproducibility and predictive capacity are formally separated. Potentially facilitating retrospective validation exercises, the main advantage of this separation, however, effects prospective validation by opening up opportunities for reduced and thus more time- and cost-efficient study designs. Taking the previous ECVAM validation study on *in vitro* methods for skin corrosivity as an example of a successful validation study – two of its methods resulted in OECD test guidelines adopted in 2004 – we analysed the feasibility of this separation. Study designs reducing the number of tests to be performed by up to 50% were simulated with the original validation data of the EPISKIN model. According to these designs, the data were re-sampled at least 10,000 times/design either randomly or stratified randomly, i.e. accounting for the potency *in vivo* in the chemical selection for reproducibility testing. We demonstrate the effects of the lean designs on the variability of several between-laboratory reproducibility measures and on the predictive capacities, in terms of sensitivity and specificity, in comparison to the original study. Overall, the study results were only little affected by the modelled lean designs. We conclude that the separation of the two modules is a promising way to speed-up prospective validation studies and to substantially reduce their costs without compromising study quality.

Lecture

Catch-up validation: Principles and case studies

Manfred Liebsch
Federal Institute for Risk Assessment, Centre for Alternative Methods to Animal Experiments – ZEBET, Berlin, Germany

In several validation studies performed on *in vitro* alternative methods, experience has shown that developers of new test procedures tried to define their procedures very tightly in order to assure (i) the best possible outcome for their test systems in these studies, (ii) to be "*primus inter pares*" in case similar test systems were available, and (iii) to gain the recognition they seemed to deserve because of their intellectual and financial investment in a new technology. However, this "artificial competition" caused significant problems in validation trials of several methods, in particular those employing reconstructed human epidermis/full skin models.

In 1994, Advanced Tissue Sciences (ATS, USA) tried to address regulatory needs of the dangerous good transport system by proposing a skin corrosion assay with the human reconstituted skin model Skin[2] that was "tuned" incredibly insensitive, just to be able to discriminate three different corrosivity transport classes. Almost expectedly, the protocol failed in the formal ECVAM skin corrosion validation study because of low sensitivity. At the same time, we had made the experience in the area of phototoxicity testing, that the Skin[2] model did not differ with regard to sensitivity/specificity, if we applied an identical protocol to different skin/epidermis models. As a consequence, when the epidermal model EPISKIN (SADUC, France) in 1998 had performed very well in the ECVAM skin corrosion validation study, but then became temporarily unavailable, ECVAM and ZEBET successfully collaborated on the concept of a "catch-up" validation study. We used the model EpiDerm (MatTek, USA) with a similar protocol and prediction model. This study opened the door to a general use of reconstructed skin/epidermal models for skin corrosion testing provided they meet structural and performance requirements defined in OECD Test Guideline 431.

Examples of currently finalised and ongoing catch-up validation studies employing other human skin models will be presented. The necessary balance between formal requirements for such studies and investment of resources when an increasing knowledge has proved high similarity of the models will be addressed.

Lecture

The ICCVAM nomination and submission process and guidelines for new, revised, and alternative test methods

Leonard M. Schechtman [1], Marilyn L. Wind [2] and William S. Stokes [3]

[1] U. S. Food and Drug Administration, National Center for Toxicological Research, Rockville, Maryland, USA; [2] U. S. Consumer Product Safety Commission, Directorate for Health Sciences, Bethesda, Maryland, USA; [3] National Institute of Environmental Health Sciences, Interagency Center for the Evaluation of Alternative Toxicological Methods, Research Triangle Park, NC, USA

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) has developed and implemented a process for the nomination and submission of test methods and for their prioritisation for review and evaluation. Prioritisation of proposed test methods is a function of their regulatory applicability, anticipated multi-agency interest and use, responsiveness to the replacement, reduction, and refinement of animal use, potential for improved predictivity of adverse effects relative to currently employed methods, and efficiency and economic savings. The newly revised ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods (http://iccvam.niehs.nih.gov/docs/guidelines/sub-guide.htm) were developed to assist test method sponsors/nominators in organising the information needed to assess the validation status of test methods at any stage of the validation process and the extent to which the ICCVAM validation and acceptance criteria have been or will be addressed. The original guidelines, in use since 1998 to evaluate the scientific validity of test methods that have since achieved regulatory acceptance, have been updated to reflect experience gained and help to facilitate a more efficient process. Adherence to these revised guidelines will help ensure the sufficiency of data and information for independent peer review and for regulatory authorities to determine the scientific validity and regulatory acceptability of test methods. The elements comprising these guidelines have now been incorporated into international guidance for the evaluation of methods proposed for new test guidelines. The ICCVAM nomination, submission, and prioritisation process and the content and organisation of submissions or nominations will be described.

Lecture

Gold standard or FeS? Are non-validated reference tests obstructing the transition to non-animal approaches in toxicology?

Troy Seidle

People for the Ethical Treatment of Animals, Research and Investigations Dept., Toronto, Canada

Modern validation standards stipulate that the accuracy and reliability of a new or revised toxicological test method should be evaluated relative to the study it is intended to replace. This proviso seems to take for granted that toxicity tests in widespread use today are valid themselves, even though only a handful of animal tests have ever been subjected to formal or rigorous validation according to ECVAM/ICCVAM/OECD criteria. Nonetheless, some insist that these tests have been "validated by convention", and regard the data generated by these tests as the "gold standard" to be met by any prospective alternative method. The integrity of this perspective will be examined relative to published Draize and developmental toxicity data, which reveal high levels of intra- and inter-laboratory variability as well as marked species differences in chemical sensitivity. These factors have been closely linked to earlier, unsuccessful efforts to validate replacements to the Draize eye irritation test. Alternate sources of reference data (e.g., occupational exposure and biomonitoring, human clinical drug trials, poison control center data, etc.) will be identified, as will the strengths and limitations of each. This analysis will provide additional support for a recommendation that emerged from an OECD validation conference in 2002: That there is a pressing need for an international workshop on the acquisition and use of human data to better evaluate the relevance and accuracy of new, revised, and existing toxicological test methods relative to the species of regulatory interest.

Lecture

The use of test method performance standards to streamline the validation process

William Stokes [1], Leonard Schechtman [2], Marilyn Wind [3], Amy Rispin [4], Kalaish Gupta [5] and Raymond R. Tice [6]

[1] NICEATM, NIEHS, NIH, Research Triangle Park, USA; [2] FDA, USA; [3] CPSC, USA; [4] EPA, USA; [5] CPSC, USA; [6] ILS, Inc., USA

Regulatory authorities are often required to communicate the basis on which new test methods have been determined to have sufficient accuracy and reliability for specific testing purposes. The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recently developed the concept of test method performance standards (PS) to address this need. PS are based on adequately validated proprietary and non-proprietary test methods that have been accepted by one or more regulatory agencies. PS can then be used to evaluate the performance of other mechanistically and functionally similar test methods that measure or predict the same biological or toxic effect. PS consist of three aspects: 1) essential test method components, which are the essential structural, functional, and procedural elements of a validated test method that should be included in the protocol of a proposed similar test method); 2) a minimum list of reference chemicals selected from the chemicals used to demonstrate acceptable performance of the validated test method, which is used to assess the accuracy and reliability of a proposed similar test method; and 3) the accuracy and reliability values that should be achieved or exceeded by the proposed test method when evaluated using the minimum list of reference chemicals. Proposed PS are developed and undergo concurrent independent peer review during the technical evaluation of a test method. The development and use of PS is expected to significantly streamline the validation and acceptance process for test methods that are mechanistically and functionally similar to accepted test methods.