

Chapter 2: The Principles and Procedures of Validation

Introduction

The aim of this chapter is to give a brief overview of the principles of validation, and of the manner in which they are applied by ECVAM, with reference to the main publications in the field. A short review by Balls & Fentem (1) explains the key concepts of validation and regulatory acceptance, whereas a consolidated account based on the many papers that describe ECVAM's principles and procedures is in preparation.

Alternative Methods, Prediction Models and Validation

An *alternative method* is any method that can be used to replace, reduce or refine the use of animal experiments in biomedical research, testing or education. The “alternatives” concept is attributed to Russell & Burch (2), who defined three types of alternatives: *reduction alternatives*, which obtain a comparable level of information from the use of fewer animals, or more information from the same number of animals; *refinement alternatives*, which alleviate or minimise potential pain, suffering and distress; and *replacement alternatives*, which permit a given purpose to be achieved without using living vertebrate animals. The three types of alternative procedure are referred to collectively as the “Three Rs”.

The *prediction model* (PM) plays an important role in the validation process. As described by Archer *et al.* (3), an alternative method for the replacement (or partial replacement) of an animal test can be thought of as the combination of a test system and a PM. The test system provides a means of generating physicochemical or *in vitro* data for chemicals of interest, whereas the PM is an unambiguous algorithm for converting these data into predictions of a pharmacotoxicological endpoint in animals or humans. The role of the PM in validation was discussed earlier by Bruner *et al.* (4), who also defined criteria for the adequacy of PMs, and described the use of computer simulations, based on the PM concept, for judging the performance of alternative tests. Subsequently, there were a number of editorials in *ATLA*, which debated the use of PMs in validation studies (5–7). A consensus view on the use of PMs was described in a joint report by the ECVAM task forces on validation and on biostatistics (3). More recently, a consolidated description of the use of PMs, including their incorporation into tiered

testing strategies, has been published from ECVAM (8).

The *validation* of an alternative method is the process by which the relevance and reliability of the method are established for a particular purpose (9). In the context of a replacement test method, *relevance* refers to the scientific basis of the test system, and to the predictive capacity of an associated PM, whereas *reliability* refers to the reproducibility of test results, both within and between laboratories, and over time (Figure 2.1). The *purpose* of an alternative method refers to its intended application, such as the regulatory testing of chemicals for a specific toxicological endpoint, for example, skin corrosivity. In other words, to validate (i.e. to establish the scientific validity of) an alternative test, it is necessary to demonstrate that for its stated purpose:

1. the test system has a sound scientific basis;
2. the predictions made by the PM are sufficiently accurate; and
3. the results generated by the test system are sufficiently reproducible within and between laboratories, and over time.

These conditions can be referred to as the criteria for scientific relevance, predictive relevance, and reliability, respectively.

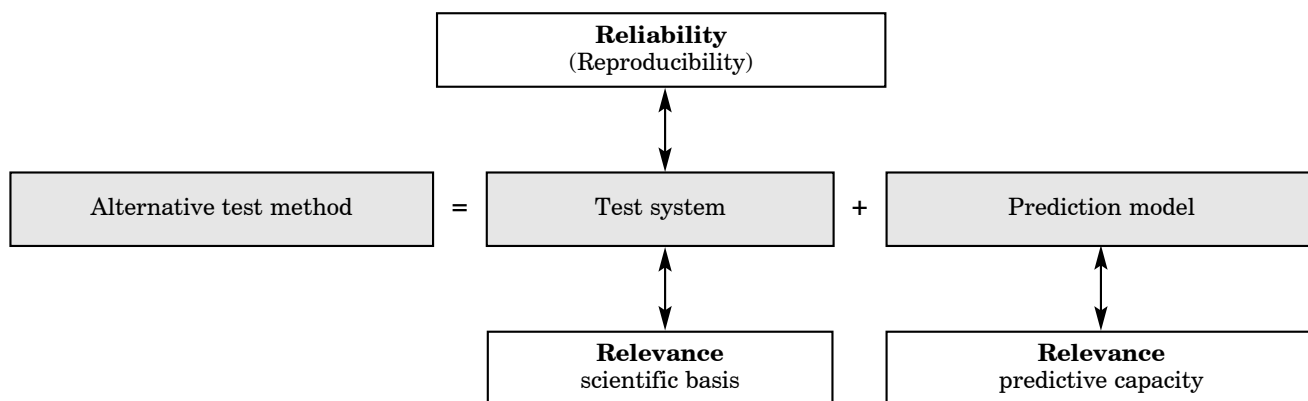
The Evolution of Alternative Methods and the ECVAM Validation Process

Five main stages in the evolution of new test methods have been identified (10, 11):

1. test development;
2. prevalidation;
3. formal validation;
4. independent assessment; and
5. progression toward regulatory acceptance.

The main role of ECVAM is to coordinate, at the European Union level, the validation of alternative methods, although ECVAM also plays a role in promoting test development and encouraging

Figure 2.1: A schematic representation of an alternative test and its performance properties



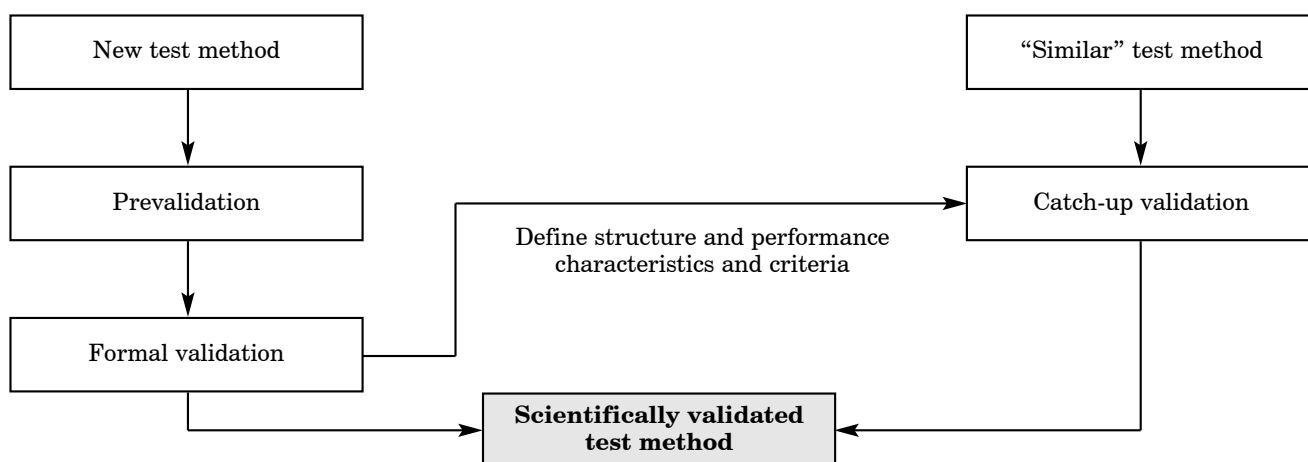
the progression of scientifically valid detailed methods toward regulatory acceptance. The roles of ECVAM, and of the ECVAM Scientific Advisory Committee (ESAC) have been described elsewhere (12).

The ECVAM validation process consists of two main stages: prevalidation and formal validation.

A *prevalidation study* is a small-scale interlaboratory study designed to refine the protocol and PM of a test method, and to obtain a preliminary assessment of its relevance and reliability. The prevalidation process is divided into three consecutive phases (13): method refinement (phase I), method transfer

(phase II), and method performance (phase III). During phase I, the protocol and PM of a test method are refined in a single laboratory (with prior experience in the use of the test). During phase II, an assessment is made of the transferability of the method to a second laboratory, making any necessary refinements to the protocol and PM. During phase III, the relevance and reliability of the test are assessed under blind conditions in three or more laboratories (which generally include the first two laboratories). An important outcome from a prevalidation study is that an optimised protocol is identified that could be used in a formal validation study.

Figure 2.2: A schematic representation of the ECVAM validation process



In the ECVAM process, a scientifically validated method is one that has been endorsed by the ECVAM Scientific Advisory Committee (ESAC). If the method is appropriate for chemicals testing, a draft Annex V guideline, incorporating the method, will be submitted to the EU Competent Authorities for Directive 67/548/EEC, for consideration for regulatory acceptance and application.

A *formal validation study* is a larger-scale inter-laboratory study, performed under blind conditions, and designed to obtain a more definitive assessment of relevance and reliability. A formal validation study can be thought of as a larger-scale version of the phase III stage of prevalidation, in which a larger number of chemicals are tested (although not necessarily in a larger number of laboratories). In general, newly developed test methods enter the prevalidation process and, following successful prevalidation, proceed to formal validation. In some cases, however, a new test may be sufficiently “similar” (in terms of its structural characteristics and reliability) to another test, the scientific validity of which has already been established, for a prevalidation study to be sufficient to establish the validity of the new method (14). This process, called “catch-up” validation, has been applied to the EpiDerm™ human skin model for skin corrosivity (15), following the successful formal validation of the EPISKIN™ human skin model (16).

A schematic showing the ECVAM validation process is given in Figure 2.2.

ECVAM’s criteria for development and validation have been defined in a number of documents (1, 10, 11, 17, 18), and are summarised in the next section. There are thought to be no major differences between these criteria and the criteria subsequently defined by ICCVAM (19, 20) and the OECD (21).

Following the completion of an ECVAM prevalidation or validation study, a report on the outcome of the study is submitted to ECVAM, so that it can be communicated to, and considered by, the ESAC. If the ESAC is satisfied that ECVAM’s criteria for test development and validation have been satisfied, an ESAC statement endorsing the scientific validity of the method is issued.

The ESAC also considers the scientific validity of methods that have been assessed by a weight-of-evidence consideration of existing information (retrospective validation), rather than by the conduct of new practical work in prevalidation and/or validation studies (prospective validation). The ESAC endorsement process, and the subsequent steps by which validated methods are submitted for regulatory consideration at the EU and OECD levels, have been described in detail elsewhere (12).

ECVAM’s Criteria for Test Development and Validation

For a method to be considered ready to enter the validation process, the following should be available, to establish that it has been sufficiently well developed:

1. A definition of the scientific purpose of the method, and of its proposed practical application.

2. A description of the basis of the method.
3. The case for its relevance. In the case of a test method for the prediction of an *in vivo* pharmacotoxicological endpoint, this should refer to the mechanistic relevance of the test, and to any preliminary evidence supporting the predictive capacity of the test.
4. An explanation of the need for the method in relation to existing *in vivo* (animal or human) methods (reference to relevant test guidelines and legislation) and other non-animal methods.
5. An optimised protocol, including: a) any necessary standard operation procedures; b) a specification of endpoints and endpoint measurements; c) the method for deriving and expressing results; d) the interpretation of the results in terms of one or more *in vivo* pharmacotoxicological endpoints, by means of a prediction model; and e) the use of adequate controls.
6. A statement about the limitations of the test.
7. Evidence of intralaboratory reproducibility and, if available, interlaboratory transferability.

In an independent assessment of the outcome of a prevalidation or validation study, the following (validation criteria) should be considered:

1. Clarity of defined goals.
2. Quality of overall design.
3. Independence of management.
4. Independence of selection, coding and distribution of test materials.
5. Independence of data collection and analysis.
6. Number and properties of test materials.
7. Quality and interpretation of results
8. Performance of the method(s) in relation to the predetermined goals of the study.
9. Reporting of outcome in the peer-review literature.
10. Availability of raw data.
11. Independence of assessment of outcome.

Practical and Logistical Aspects of Prevalidation and Validation Studies

The design, management and conduct of prevalidation and validation have been described in a number of publications. Curren *et al.* (13) focused on the role of prevalidation, describing in detail the three phases of this process. Balls *et al.* (10) described: a) the role of the Management Team in a validation study; b) the selection of alternative tests, laboratories and chemicals; c) compliance with safety standards and the principles of Good Laboratory Practice (GLP); d) the collection and analysis of experimental data; and e) the assessment of the outcome of the validation studies. Problems encountered in some early validation studies, difficulties associated with the selection and use of *in vitro* data, and the comparison of *in vitro* and *in vivo* endpoints, were discussed by Balls & Fentem (1).

Recommendations concerning the application of GLP principles to *in vitro* toxicology, and in particular, to validation studies, have been published by Cooper-Hannan *et al.* (22).

Recommendations for the application of statistical methods in the validation process were made by Holzhütter *et al.* (23), and there are several examples of the use of biostatistical methods for the development and assessment of PMs (24–26).

Examples of Prevalidation and Validation Studies

Reports on a number of ECVAM prevalidation and validation studies have been published (for example, 27, 28). In terms of their compliance with currently accepted principles and procedures, the ECVAM skin irritation prevalidation study (29) can be regarded as a model for an ECVAM prevalidation study, whereas the ECVAM skin corrosivity validation study (16) and the EU/COLIPA phototoxicity validation study (30) can be regarded as model validation studies. The procedures involved in the selection of test chemicals for a validation study are well illustrated by Barratt *et al.* (31).

As an example, the ECVAM skin corrosivity validation study (16, 31) was conducted as a follow-up to an ECVAM prevalidation study (27). The main objectives of the validation study were to: a) identify tests capable of discriminating corrosives from non-corrosives for selected types of chemicals and/or for all chemicals; and b) determine whether these tests could correctly identify known R35 and R34 chemicals. The tests evaluated were the rat skin transcutaneous electrical resistance (TER) assay, CORROSITEX™, the Skin²™ ZK1350 corrosivity test and EPISKIN™. Each test was conducted in three independent laboratories (to assess

interlaboratory transferability), on at least two occasions (to assess interlaboratory reproducibility), with 60 coded chemicals. Since three laboratories were used to perform experimental work on each of four tests, a total of 12 laboratories were involved. Some of these laboratories were based in the EU and some in the USA, showing that validation studies are international collaborative exercises.

All the tests evaluated showed acceptable intralaboratory and interlaboratory reproducibilities. The TER, Skin² and EPISKIN tests proved to be applicable for testing a diverse group of chemicals of different physical forms, including organic acids, organic bases, neutral organics, inorganic acids, inorganic bases, inorganic salts, electrophiles, phenols and soaps/surfactants. Two of the four tests evaluated, the TER assay and EPISKIN, met the criteria agreed in advance by the Management Team concerning acceptable underprediction and overprediction rates, so they could be considered scientifically validated for use as replacements for the animal test for distinguishing between corrosive and non-corrosive chemicals for all of the chemical types studied. EPISKIN was the only test able to distinguish between known R35 and R34 chemicals, for all of the chemical types included, on an acceptable number of occasions. The corrosive potentials of about 40% of the test chemicals could not be assessed with CORROSITEX, and the assay did not meet the criteria necessary for it to be acceptable as a replacement test. However, CORROSITEX was considered to be valid for testing specific classes of chemicals, such as organic bases and inorganic acids. The Skin² assay did not meet the criteria required for it to be considered scientifically validated. On the basis of the successful outcome of this validation study, the ESAC subsequently endorsed the use of the EPISKIN (32) and TER (33) methods.

The Time Required for Validation and Regulatory Acceptance

There is widespread dissatisfaction at the rate of adoption of validated methods into regulatory requirements and testing guidelines. As an illustration, Table 2.1 gives the timing of the most important steps leading from the ECVAM prevalidation study on *in vitro* methods for skin corrosivity, to the validation study and the regulatory acceptance of two skin corrosivity methods at the EU and OECD levels.

The rate-limiting steps in the evolution of alternative methods appear to be the rate at which suitably developed methods become available for entry into the prevalidation process, and the rate at which scientifically validated methods are accepted by regulatory authorities. The

Table 2.1: Main steps in the validation and regulatory acceptance of *in vitro* methods for skin corrosivity

Activity	Date	
1. Prevalidation study		1993–1994
2. Meeting to discuss proposed validation study	8–9 February	1995
3. Publication of outcome of prevalidation study in <i>ATLA</i>	April	1995
4. Chemical selection meeting	19 September	1995
5. Publication of background information on study	November	1995
6. 1st Management Team meeting	11 January	1996
7. Confirmation of participating laboratories	March	1996
8. Meeting to select final set of 60 test chemicals	15 April	1996
9. 2nd Management Team meeting	30 April	1996
10. Refinement of test protocols	May	1996
11. Training of laboratory personnel	May	1996
12. Distribution of first set of 10 coded chemicals	June	1996
13. Definition of data-analysis procedures	August	1996
14. Preliminary data analysis (first 10 chemicals)	September	1996
15. 3rd Management Team meeting	16 September	1996
16. Distribution of remaining 50 chemicals	September	1996
17. Submission of results for second set of 50 chemicals	March	1997
18. Preliminary data analysis for all 60 test chemicals	June	1997
19. 4th Management Team meeting	18–20 June	1997
20. Statement on study outcome issued by Management Team	18 July	1997
21. 5th Management Team meeting	25 September	1997
22. Submission of reports for publication	October	1997
23. Reports accepted for publication in <i>Toxicology in Vitro</i>	April	1998
24. Endorsement of EPISKIN and TER by ESAC	April	1998
25. Report published in <i>Toxicology in Vitro</i>	August	1998
26. Draft EU Test Method submitted to EU National Coordinators	September	1998
27. Draft OECD Test Guideline submitted to OECD	September	1998
28. Acceptance of Test Method on skin corrosion by EU regulators	4 February	2000
29. Test Method on skin corrosion incorporated into Annex V of <i>Directive 67/548/EEC</i> by means of Commission <i>Directive 2000/33/EC</i>	25 April	2000
30. OECD Extended Expert Consultation on the draft OECD Test Guideline	1–2 November	2001
31. OECD circulates two new draft OECD Test Guidelines: TG 430 (TER) and TG 431 (human skin model) for comment	27 March	2002
32. OECD National Coordinators approve TG 430 and TG 431	31 May	2002

ESAC = ECVAM Scientific Advisory Committee

TER = transcutaneous epidermal resistance

“ECVAM-dependent” part of the process, which generally comprises both the prevalidation and validation steps, typically takes 4–6 years. However, adoption of the fast-track validation procedure, catch-up validation (15, 34), for appropriate methods, can reduce the time taken to less than two years.

References

- Balls, M. & Fentem, J.H. (1999). The validation and acceptance of alternatives to animal testing. *Toxicology in Vitro* **13**, 837–846.
- Russell, W.M.S. & Burch, R.L. (1959). *The Principles of Humane Experimental Technique*, 238pp. London, UK: Methuen.
- Archer, G., Balls, M., Bruner, L.H., Curren, R.D., Fentem, J.H., Holzhütter, H.-G., Liebsch, M., Lovell, D.P. & Southee, J.A. (1997). The validation of toxicological prediction models. *ATLA* **25**, 505–516.
- Bruner, L.H., Carr, G.J., Chamberlain, M. & Curren, R.D. (1996). Validation of alternative methods for toxicity testing. *Toxicology in Vitro* **10**, 479–501.
- Bruner, L., Carr, G., Chamberlain, M. & Curren, R. (1996). No prediction model — no validation study. *ATLA* **24**, 211–217.
- Archer, G. (1996). No statistics, no science — let

- alone validation studies. *ATLA* **24**, 453–456.
7. Holzhütter, H-G. & Spielmann, H. (1996). The use of prediction models in toxicology: what should they predict? *ATLA* **24**, 881–884.
 8. Worth, A.P. & Balls, M. (2001). The importance of the prediction model in the development and validation of alternative tests. *ATLA* **29**, 135–143.
 9. Balls, M., Blaauboer, B., Brusick D., Frazier J., Lamb, D., Pemberton, M., Reinhardt, C., Roberfroid, M., Rosenkranz, H., Schmid, B., Spielmann, H., Stamatii, A-L. & Walum, E. (1990). Report and recommendations of the CAAT/ERGATT workshop on the validation of toxicity test procedures. *ATLA* **18**, 313–337.
 10. Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H. & Zucco, F. (1995). Practical aspects of the validation of toxicity test procedures. The report and recommendations of ECVAM workshop 5. *ATLA* **23**, 129–147.
 11. Fentem, J.H. & Balls, M. (1997). The ECVAM approach to validation. In *Developments in Animal and Veterinary Sciences, Volume 27, Animal Alternatives, Welfare and Ethics. Proceedings of the 2nd World Congress on Alternatives and Animal Use in the Life Sciences, Utrecht, The Netherlands, 20–24 October 1996* (ed. L.F.M. van Zutphen & M. Balls), pp. 1083–1089. Amsterdam, The Netherlands: Elsevier.
 12. Worth, A.P. & Balls, M. (2001). The role of ECVAM in promoting the regulatory acceptance of alternative methods in the European Union. *ATLA* **29**, 525–535.
 13. Curren, R.D., Southee, J.A., Spielmann, H., Liebsch, M., Fentem, J.H. & Balls, M. (1995). The role of prevalidation in the development, validation and acceptance of alternative methods. *ATLA* **23**, 211–217.
 14. Balls, M. (1997). Defined structural and performance criteria would facilitate the validation and acceptance of alternative test procedures. *ATLA* **25**, 483–484.
 15. Liebsch, M., Traue, D., Barrabas, C., Spielmann, H., Uphill, P., Wilkins, S., McPherson, J.P., Wiemann, C., Kaufmann, Remmele, M. & Holzhütter, H-G. (2000). The ECVAM prevalidation study on the use of EpiDerm for skin corrosivity testing. *ATLA* **28**, 371–401.
 16. Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhütter, H.G. & Liebsch, M. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. II. Results and evaluation by the management team. *Toxicology in Vitro* **12**, 483–524.
 17. Balls, M. & Karcher, W. (1995). The validation of alternative test methods. *ATLA* **23**, 884–886.
 18. Balls, M. & Fentem, J.H. (1997). Progress toward the validation of alternative tests. *ATLA* **25**, 33–43.
 19. NIH (1997). *Validation and Regulatory Acceptance of Toxicological Test Methods. A Report of the Ad Hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Publication 97–3981*, 105 pp. Research Triangle Park, NC, USA: NIEHS.
 20. NIH (1999). *Evaluation of the Validation Status of Toxicological Methods: General Guidelines for Submissions to ICCVAM (Revised, October 1999). NIH Publication 99–4496*, 44pp. Research Triangle Park, NC, USA: NIEHS.
 21. OECD (1996). *Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods*, 60 pp. Paris, France: OECD.
 22. Cooper-Hannan, R., Harbell, J.W., Coecke, S., Balls, M., Bowe, G., Cervinka, M., Clothier, R., Hermann, F., Klahm, L.K., de Lange, J., Liebsch, M. & Vanparys, P. (1999). The principles of Good Laboratory Practice: application to *in vitro* toxicology studies. The report and recommendations of ECVAM workshop 37. *ATLA* **27**, 539–577.
 23. Holzhütter, H-G., Archer, G., Dami, N., Lovell, D.P., Saltelli, A. & Sjöström, M. (1996). Recommendations for the application of biostatistical methods during the development and validation of alternative toxicological methods. ECVAM biostatistics task force report 1. *ATLA* **24**, 511–530.
 24. Spielmann, H., Liebsch, M., Kalweit, S., Moldenhauer, F., Wirnsberger, T., Holzhütter, H-G., Schneider, B., Glaser, S., Gerner, I., Pape, W.J.W., Kreiling, R., Krauser, K., Miltenburger, H.G., Steiling, W., Luepke, N.P., Muller, N., Kreuzer, H., Murmann, P., Spengler, J., Bertram-Neis, E., Siegemund, B. & Wiebel, F.J. (1996). Results of a validation study in Germany on two *in vitro* alternatives to the Draize eye irritation test, the HET-CAM test and the 3T3-NRU cytotoxicity test. *ATLA* **24**, 741–858.
 25. Genschow, E., Scholz, G., Brown, N.A., Piersma, A.H., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S. & Spielmann, H. (2000). Development of prediction models for three *in vitro* embryotoxicity tests in an ECVAM validation study. *In Vitro and Molecular Toxicology* **13**, 51–65.
 26. Worth, A.P. & Cronin, M.T.D. (2001). Prediction models for eye irritation potential based on endpoints of the HET-CAM and neutral red assays. *In Vitro and Molecular Toxicology* **14**, 143–156.
 27. Balls, M., Botham, P.A., Bruner, L.H. & Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology in Vitro* **9**, 871–929.
 28. Botham, P.A., Chamberlain, M., Barratt, M.D., Curren, R.D., Esdaile, D.J., Gardner, J.R., Gordon, V.C., Hildebrand, B., Lewis, R.W., Liebsch, M., Logemann, P., Osborne, R., Ponec, M., Régnier, J.F., Steiling, W., Walker, A.P. & Balls, M. (1995). A prevalidation study on *in vitro* skin corrosivity testing. The report and recommendations of ECVAM workshop 6. *ATLA* **23**, 219–255.
 29. Fentem, J. H., Briggs, D., Chesné, C., Elliott, G. R., Harbell, J. W., Heylings, J. R., Portes, P., Roguet, R., van de Sandt, J. J. M. & Botham, P. A. (2001). A prevalidation study on *in vitro* tests for acute skin irritation: results and evaluation by the Management Team. *Toxicology in Vitro* **15**, 57–93.
 30. Spielmann, H., Balls, M., Dupuis, J., Pape, W.J., Pechovitch, G., de Silva, O., Holzhütter, H.G., Clothier, R., Desolle, P., Gerberick, F., Liebsch, M., Lovell, W.W., Maurer, T., Pfannenbecker, U., Potthast, J.M., Csato, M., Sladowski, D., Steiling, W. & Brantom, P. (1998). The International EU/COLIPA *In Vitro* Phototoxicity Validation Study: Results of Phase II (Blind Trial). Part 1: The 3T3

- NRU Phototoxicity Test. *Toxicology In Vitro* **12**, 305–327.
31. Barratt, M.D., Brantom, P.G., Fentem, J.H., Gerner, I., Walker, A.P. & Worth, A.P. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 1. Selection and distribution of the test chemicals. *Toxicology in Vitro* **12**, 471–482.
 32. ECVAM (1998). Statement on the scientific validity of the EPISKIN™ test (an *in vitro* test for skin corrosivity). *ATLA* **26**, 277–280.
 33. ECVAM (1998). Statement on the scientific validity of the rat skin transcutaneous electrical resistance (TER) test (an *in vitro* test for skin corrosivity). *ATLA* **26**, 275–277.
 34. ECVAM (2001). Statement on the application of the EPIDERM® human skin model for skin corrosivity testing. *ATLA* **28**, 365–367.