



Food for Thought ... on Alternative Methods for Cosmetics Safety Testing

Thomas Hartung

EU JRC, TRiVA unit, Ispra, Italy

This is the first article of this series published after I changed responsibility (see *ALTEX* 2, 2008). Thus, I considered whether it is appropriate to continue this series. Since the contributions are written in a clearly personal way and a (thought) provocative style, they can hardly be mistaken as official positions of the European Commission. Most probably, half of the ideas expressed are wrong anyway (but nobody knows which half). If you consider, however, that about 30% of scientific articles are not reproducible and that the scientific publishing culture usually does not permit the putting forward of concepts and ideas without experimental proof, this difference in “proportion of truth” might be compensated for the reader. Together with the editor of *ALTEX*, I decided to continue this series – I have been working in the field for more than 20 years now and consider my views as those of an individual scientist imprinted by the experiences of the last few years. However, I will refrain from any discussions of ECVAM’s future work as a validation body. I appreciated very much that my predecessor at ECVAM, Michael Balls, at the time gave me the opportunity to give things “my spin” without any interference, and once again I will try to follow his role model here.

The 7th amendment (directive 2003/15/EC) of the cosmetics directive (*Directive 76/768/EEC*) was my “welcome present” when I started at ECVAM in late 2002. Beside REACH, which started only just now on 1st of June 2008 with the agency taking over, this amendment has shaped the landscape of alternative methods dramatically. It represented a prime opportunity to give alternative methods a new, enlarged role and the incentive for support by a major industry.

Recital 5 of the directive reads: “*Currently, only alternative methods which are scientifically validated by the European Centre for the Validation of Alternative Methods (ECVAM) or the Organisation for Economic Cooperation and Development (OECD) and applicable to the whole chemical sector are systematically adopted at Community level. However, the safety of cosmetic products and their ingredients may be ensured through the use of alternative methods which are not necessarily applicable to all uses of chemical ingredients. Therefore, the use of such methods by the whole cosmetic industry should be promoted and their adoption at Community level ensured, when such methods offer an equivalent level of protection to consumers.*” This clearly indicates that the legislator wants to uncouple this industry from the possibly slower progress in other areas of application of alternative methods.

Key features of the directive with regard to alternative approaches are the deadlines for phasing out testing and their reinforcement by marketing bans. To recapitulate, there are four deadlines:

- A deadline of 11 September 2004 for the testing of finished products (reinforced by a marketing ban)
- An immediate testing ban for ingredients if an alternative method is “*validated and adopted at Community level with due regard to the development of validation within the OECD*”. This leaves some room for interpretation. Depending on the view, this means after the validity statement by ECVAM’s Scientific Advisory Committee, after this “advice” is taken and adopted by ECVAM and the European Commission, after its acceptance by DG SANCO’s Scientific Committee on Consumer Products (SC-

CP) or acceptance and inclusion in the EU test guideline regulation (formerly Annex V of the Dangerous Substance Directive, currently transformed into an independent legislation) or even acceptance as an OECD test guideline.

- A general testing ban on cosmetic ingredients from 11 March 2009, reinforced for 10 animal test requirements by an instant marketing ban.
- A marketing ban from 11 March 2013 for the more complex endpoints (those requiring repeated substance application, e.g. repeat dose toxicity, sensitization, reproductive toxicity and carcinogenicity as well as toxicokinetics, which is actually no typical testing demand). Noteworthy, the legislation does foresee a review of the feasibility of the 2013 deadline in 2011 and can further postpone this in a co-decision procedure.

The legislation is in many ways unique as it, for example, phases out essential safety tests before alternatives are available. This “incentive for change” probably reflects the legislators’ discontent that the 6th amendment of 1993 led only to two postponements of the already foreseen phasing out of animal testing. The 6th amendment introduced a marketing ban on cosmetic products tested on animals from 1 January 1998, provided that alternative testing methods had been validated and accepted by that date. The marketing ban has been postponed twice by the EU, on the grounds that insufficient progress had been made in developing and validating alternatives to the animal tests used for assessing cosmetic safety. Noteworthy, the 7th amendment permits from 2009/2013 onwards only the use of replacement alternatives.



Some thoughts a few months before the most critical deadline of 2009:

Consideration 1: Cosmetic industry is the wrong victim, but we can be happy to have one

Animal numbers used for cosmetics are very low: The official statistics for the EU list less than three thousand for 2002 and less than six thousand for 2005 (all in France!), which represents 0.05% of all laboratory animal use. According to the cosmetic industry, no finished cosmetic product has been tested on animals since 1989. So what's all the fuss? These numbers, however, contrast remarkably with the exchange of products in this industry: This European industry represents 2,000 relatively profitable companies with 60 billion € turnover. The sector is characterised by quick product exchange (5,000 new products in Europe and 22,000 world-wide per year, 25% of turnover with products released within the last 6 months). Market leader L'Oréal, for example, releases 3,000 new products per year and out of 500 patents about 100 are patents on substances. It has to be assumed that several hundred new substances are introduced into cosmetics every year. Given about 8,000 cosmetic ingredients in use in total, this number represents a reasonable assumption of turnover. This might be compared with only 8 new active substances entering the pharmaceutical world-market on average per year. Still, testing cosmetic products on animals is rare to negligible; although safety must be assured for products we put on our skin, into our eyes and into our mouth. The reason for rare animal testing: the cosmetic industry is not producing its chemical ingredients; ingredients are tested as chemicals and many food ingredients as well as natural products are used. Thus, testing is done by the ingredient manufacturer and not commissioned by cosmetic industry. This might, however, change now (see consideration 2).

Cosmetic industry has invested into alternatives, mainly with their SCAAT programme (COLIPA's joint industry research programme on Alternative Approaches to Animal Testing), and some

of the global companies have invested individually (e.g. L'Oréal, P&G, Unilever, etc.), focussing on the "cosmetic endpoints" of skin/eye toxicity and sensitization (allergy). However, over the last 2-3 years, substantial investments and efforts have been made related to other endpoints, turning this industry at the moment into the main driver and partner for alternative methods. Without doubt, the 7th amendment was a strong incentive for this engagement. Not least due to the 7th amendment, cosmetic industry has embraced alternative methods; in words of Dr. Raniero De Stasio, chair of the COLIPA Communications Project Team on Alternative Testing: "Alternatives also open up a whole new range of possibilities that improve the tools leading to innovation," he says. "They allow scientists in cosmetics to move forward faster." (COLIPA activity report, 2007; <http://www.colipa.com/site/index.cfm?SID=15588>)

The 7th amendment required the Commission, after consultation of the SCCNFP (now SCCP) and ECVAM, and with due regard to the development of validation within the OECD, to establish timetables for the implementation of the provisions, including deadlines for the phasing-out of the various animal tests. In view of establishing these timetables, the Commission decided to set up an "ad-hoc group" between Commission services, stakeholder representatives for industry, animal welfare and consumer associations, and the OECD. The work was based on an ECVAM report from 2002 (Worth and Balls, 2002), which had also strongly influenced the timeline setting of the legislation. The participants agreed on nominating experts for the 11 human health effects of concern in order to gain scientific expertise, and ECVAM, besides participating with its expertise in all of the endpoint working groups, also coordinated and steered the scientific process (Eskes and Zuang, 2005).

The Directive requires also that the Commission present a yearly monitor of these timetables and decide on possible adaptation of them within the maximum periods (6 or 10 years). ECVAM usually writes the technical part of these Annual Reports, which are submitted to Council and the European Parliament. This

gives the field of alternatives a remarkable political visibility, which is further enhanced by the continuous work of the European Parliament InterGroup for animal welfare, which holds a monthly meeting and involves about 40 members of parliament.

ECVAM is explicitly anchored in the legislation (hurrah!): recital 5 cited above, recital 7 "It will gradually become possible to ensure the safety of ingredients used in cosmetic products by using non-animal alternative methods validated at Community level, or approved as being scientifically validated, by ECVAM..." and article 1.2 "The Commission, after consultation of the SCCNFP and of the European Centre for the Validation of Alternative Methods (ECVAM) and with due regard to the development of validation within the OECD, shall establish timetables for the implementation of the provisions under paragraph 1(a), (b) and (d), including deadlines for the phasing out of the various tests.". This is an important safeguard for the maintenance of this service, as it is only mentioned in an Annex of the REACH legislation once and its anchoring in the revision of directive 86/609/EEC (which dates before the creation of ECVAM) is unclear.

The new chemical legislation REACH represents another incentive for introducing alternative methods, at first glance addressing all the toxicological effects relevant for cosmetics. The situation for cosmetics, however, is very different to the one of REACH: While REACH would benefit from any reduction or refinement of animal tests and could stand areas without replacements, here, with an even shorter deadline (2009), full replacement has to be achieved for some selected endpoints. In contrast, animal testing for REACH will start mostly after 2015. It is evident that the short timeline only allowed having methods validated over the last few years which were already available and did not permit new developments. The inventory of the DG ENTR/ECVAM stakeholder group formed the basis for strategy development for the most relevant areas (skin corrosion, phototoxicity, skin irritation, eye irritation, skin penetration, mutagenicity, acute toxicity and sensitization).

Tab. 1: Status of scientific work, validation activities and regulatory acceptance for the main toxicological endpoints relevant for the 2009 deadline of the 7th amendment

Toxicological endpoint	R&D	Validation	ESAC statement	EU regulatory acceptance	OECD test guideline
<i>Skin corrosion</i>		Similar tests (human skin models; currently 1)	human skin models: Episkin (1998), EpiDerm (2000), Skin Ethic (2006), EST-1000 (2008), TER (1998), 1997 Corrositex ¹	2000, pending update by 30 th ATP and in draft regulation of test methods	TG 430 and TG 431 (additional data needed for the confirmation of non-corrosives)
<i>Acute phototoxicity</i>		Integrated test strategy with 3T3 NRU-PT and human skin model to reduce "false" positive rates	1997 (3T3 NRU-PT)	2000, pending update by 30 th ATP and in draft regulation of test methods	TG 432
<i>Skin absorption/penetration</i>		Submission of dossier awaited, method already in use without formal validation		Pending through 30 th ATP and in draft regulation of test methods	TG 428 (Excised pig or human skin in flow-through or static diffusion cells)
<i>Skin irritation</i>	Follow-up: omics approach to identify new markers for better discrimination of mild irritants (GHS)	Several similar tests (currently 4)	2007 (human skin model: Episkin) ²	TG under preparation for human skin model assays	TG under preparation for human skin model assays
<i>Eye irritation</i>	Collaboration with COLIPA and ICCVAM for the development of more mechanistic assays	10 tests: 4 cell-based assays, 2 reconstituted human tissue models, 2 organotypic assays, Irritection, Slug Mucosal test	2007 on the BCOP and ICE to identify severe eye irritants ⁴ Additional work needed for the HET-CAM and IRE Statement on LVET ⁵ pending	TG under preparation for BCOP and ICE	TG under preparation for BCOP and ICE
<i>Genotoxicity/mutagenicity</i>	COMICS (STREP) Reduction of the false positives in <i>in vitro</i> tests; Omission of positive/negative controls in <i>in vivo</i> genotoxicity testing	Genotoxicity tests in 3D human skin models	2006 on the micronucleus test <i>in vitro</i> ³	MNT <i>in vitro</i> is mentioned in Annex VIII of regulation (EC) no 1907/2006 (REACH)	Finalisation of test guideline on MNT <i>in vitro</i>
<i>Acute toxicity</i>		A-cute-Tox IP Predictive capacity of 3T3 NRU cytotox. assay to classify non-toxic substances	Statement on the ATC, UDP, FDP (2007) ⁶	2001 ATC, UDP, FDP	TG 420, TG 423, TG 425

¹ Corrositex is only applicable to acids and basis and their derivatives.

² The other validated model, EpiDerm, can only be used as part of a test strategy to identify skin irritants.

³ The Micronucleus test *in vitro* is part of a test strategy for genotoxicity/mutagenicity. Replacement of the animal tests for the identification of **genotoxic** chemicals is unlikely to happen by 2009. On the other hand, the current test strategy allows the identification of **non-genotoxic** substances on the basis of *in vitro* tests only.

⁴ several *in vitro* tests are currently under evaluation at ECVAM for the identification of mild and non-irritant substances. Test strategies to combine the different validated *in vitro* tests in order to fully replace the animal test were designed.

⁵ only refinement methods

⁶ only reduction methods



Furthermore, the sponsoring of development of alternative methods in the EU Framework Program 6 with about 100 million € was much more tailored to the needs of chemicals and cosmetics safety assessment. As a result of these efforts, 171 methods (status December 2006) entered evaluation/validation; they are at very different stages and in very different programmes, however, all of the methods have a standardised protocol (i.e. a standard operating procedure as defined by the OECD guidance document on GLP and *In vitro* Methods), are considered promising and are at least undergoing reproducibility assessments. The majority is undergoing prevalidation in Framework Programme projects. 37 methods are at late stages of validation (large-scale ring trials). All methods have been developed up to a stage at which they can be considered for validation; key developers are the Framework Programme consortia and industry. 67 of the methods which are being evaluated involve ECVAM sponsoring, 89 of them receive Framework Programme funding; while only 8 studies are financed by industry alone, 15 of them are co-sponsored by ECVAM in collaboration with industry or other national validation bodies. Noteworthy, however, the industry (co)sponsored methods include many of the advanced 37 methods. Only few (7) are sponsored by US or academia.

We should not forget that the 7th amendment also commits the Commission to further alternative methods, notably also outside the EU (recital 10): “The recognition by non-member countries of alternative methods developed in the Community should be encouraged. In order to achieve this objective, the Commission and the Member States should take all appropriate steps to facilitate acceptance of such methods by the OECD. The Commission should also endeavour, within the framework of European Community cooperation agreements, to obtain recognition of the results of safety tests carried out in the Community using alternative methods so as to ensure that the export of cosmetic products for which such methods have been used is not hindered and to prevent or avoid non-member countries requiring the repetition of such tests using animals.” This

was certainly a driving force behind last year’s creation of the International Collaboration on Cosmetic Regulation between the EU, the US, Japan and Canada, which already focused strongly on alternative methods at the first meeting and encouraged international collaboration in its resolution: “ICCR recognised the importance of reducing, refining and replacing animal testing. The group welcomed the efforts of industry and validation centres in developing and validating scientific alternatives to animal testing. Intensive collaboration and communication in the design, execution, and peer review of validation studies should be further strengthened. ICCR invites IC-CVAM, ECVAM, JaCVAM and a knowledgeable representative of the Government of Canada to address this issue and to propose options to ensure a collaborative approach to this issue. They should be supported by scientific experts from the regulatory bodies.” As a direct consequence, my proposal presented at last year’s World Conference in Tokyo to create an International Council of Validation Bodies (ICVaBo, see also Bottini et al., 2007) was furthered in a series of meetings, and such collaboration was further endorsed by the April 2008 Transatlantic Economic Council: “The US Food and Drug Administration (FDA) and the European Commission have agreed to meet regularly to further their cooperation in the peer review of unvalidated alternative methods to animal testing used to determine the safety of cosmetic ingredients (including some products regulated in the United States as drugs and in the European Union as cosmetics).” (cited from White House: <http://www.whitehouse.gov/news/releases/2008/06/20080610-4.html>). This shows that legislation is starting to have an impact beyond the EU, not only in market forces but also in international political efforts.

There are a number of ongoing activities as a result of the 7th amendment. Table 1 summarises the scientific work and the completed statements on validity relevant to cosmetics are summarised in Table 2. Thus, once again: We most probably have the wrong victim (minor animal user), but we are lucky to have one.

Consideration 2: The 7th amendment becomes a threat to industry mainly due to REACH and global markets

Cosmetic industry is perhaps less threatened by the legislation as long as its ingredients are tested as chemicals. However, cosmetics ingredients were explicitly exempted from REACH. The critical question will be whether new or existing chemicals tested for other purposes on animals can still be used for cosmetics. The interpretation of advocate general Leendert Geelhoed at the European Court of Justice (ECJ) in the court case of France against the 7th amendment in May 2005 is key here (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:62003C0244:EN:HTML>): “First, it seems clear that the ban on animal tests applies equally to tests performed for the purposes of complying with other legislation, in so far as substances that have been the subject of such tests may not be used as or in cosmetic products. This interpretation seems necessary for the effect utile of the Directive and is consistent with the intention expressed in the preparatory documents leading up to its adoption.” But also his further interpretation deserves attention: “Second, it follows in my view from the wording of the contested provision that it applies to the performance of animal testing of cosmetic products or ingredients on a Member State’s territory, irrespective of whether this testing is for products destined for export. This interpretation is also suggested by Article 1(7) of Directive 2003/15. Third, it follows equally from this wording that cosmetic products and ingredients subject to animal tests outside the Community are subject to the marketing ban. Such tests would by their nature have been performed in order to meet public health requirements, thus falling within the prohibition.” This interpretation is very clear: No animal test for other countries, in other countries or under other legislations! The latter is most probably the most questionable one – if REACH is now going to readdress 30,000 old chemicals, it would be detrimental if this excludes them from further use in cosmetics. Sure, REACH explicitly exempts cosmetic ingredients, but most substances have multiple uses.

Tab. 2: ESAC statements on alternative methods relevant for cosmetics

No.	Method	Date of ESAC statement	ECVAM prospective validation study?	Impact on 3Rs	ATLA Reference
1	3T3 NRU phototoxicity test	03/11/1997	Yes	Replacement OECD TG 432	ATLA 26 (1), 7-8
2	EpiSkin skin corrosivity test	03/04/1998	Yes	Replacement (EU) OECD TG 431	ATLA 26 (3), 275-280
3	Rat TER skin corrosivity test	03/04/1998	Yes	Replacement (EU) OECD TG 430	ATLA 26 (3), 275-280
4	Application of the 3T3 NRU phototoxicity test to UV filter chemicals	20/05/1998	Yes	Replacement OECD TG 432	ATLA 26 (4), 383-386
6	Local lymph node assay for skin sensitization	21/03/2000	No	Reduction/Refinement OECD TG 429	ATLA 28 (3), 365-367
7	EpiDerm skin corrosivity test	21/03/2000	Yes	Replacement (EU) OECD TG 431; Annex V TG B.42	ATLA 28 (3), 365-367
8	CORROSITEX skin corrosivity test	06/12/2000	Yes	Reduction OECD Draft TG 435	ATLA 29 (2), 93-97
11	Micromass embryotoxicity assay	01/05/2002	Yes	Reduction	ATLA 30 (3), 265-273
12	Whole rat embryotoxicity assay	01/05/2002	Yes	Reduction	ATLA 30 (3), 265-273
13	Embryonic stem cell test for embryotoxicity	01/05/2002	Yes	Reduction OECD Draft GD 43	ATLA 30 (3), 265-273
17	Upper Threshold Concentration (UTC) step-down approach for acute aquatic toxicity testing	21/03/2006	Yes	Reduction Submitted to OECD	ATLA 35 (2), 199-208
18	CFU-GM assay for predicting acute neutropenia in humans	21/03/2006	Yes	Reduction Submitted to EMEA	ATLA 35 (2), 199-208
19	Human Whole Blood IL-1 for <i>in vitro</i> pyrogenicity testing	21/03/2006	Yes	Replacement Submitted to EMEA and European Pharmacopoeia; drafting of a general monograph is in progress	ATLA 35 (2), 199-208
20	Human Whole Blood IL-6 for <i>in vitro</i> pyrogenicity testing	21/03/2006	Yes	Replacement Submitted to EMEA and European Pharmacopoeia; drafting of a general monograph is in progress	ATLA 35 (2), 199-208
21	PBMC IL-6 for <i>in vitro</i> pyrogenicity testing	21/03/2006	Yes	Replacement Submitted to EMEA and European Pharmacopoeia; drafting of a general monograph is in progress	ATLA 35 (2), 199-208
22	MM6 IL-6 for <i>in vitro</i> pyrogenicity testing	21/03/2006	Yes	Replacement Submitted to EMEA and European Pharmacopoeia; drafting of a general monograph is in progress	ATLA 35 (2), 199-208
23	Human Cryopreserved Whole Blood IL-1 for <i>in vitro</i> pyrogenicity testing	21/03/2006	Yes	Replacement Submitted to EMEA and European Pharmacopoeia; drafting of a general monograph is in progress	ATLA 35 (2), 199-208
24	<i>In vitro</i> micronucleus test as an alternative to the <i>in vitro</i> chromosome aberration assay for genotoxicity testing	17/11/2006	Yes	Enhancement of <i>in vitro</i> test battery, OECD Draft Guideline 827	ATLA 35 (2), 199-208
25	Application of the SkinEthic human skin model for skin corrosivity testing	17/11/2006	No	Replacement (EU) OECD TG 431	ATLA 35 (2), 199-208
27	Bovine Corneal Opacity and Permeability (BCOP) test method	27/04/2007	No	Reduction Will be submitted to OECD via EC	ATLA 35 (3), 303-312
28	Isolated Chicken Eye (ICE) test method	27/04/2007	No	Reduction Will be submitted to OECD via EC	ATLA 35 (3), 303-312
29	Reduced Local Lymph Node Assay (rLLNA)	27/04/2007	Yes	Reduction OECD submission	ATLA 35 (3), 303-312
30	EpiDerm (with MTT reduction) for skin irritation	27/04/2007	Yes	Reduction	ATLA 35 (3), 303-312
31	EPISKIN (with MTT reduction) for skin irritation	27/04/2007	Yes	Replacement Will be submitted to OECD via EC	ATLA 35 (3), 303-312
32	Fixed dose procedure (FDP)	31/10/2007	No	OECD TG 420	ATLA 36(1), 12-13
33	Acute Toxic Class Method (ATC)	31/10/2007	No	OECD TG 423	ATLA 36(1), 12-13
34	Up and Down procedure (UDP)	31/10/2007	No	OECD TG 425	ATLA 36(1), 12-13
35	EST-1000 for skin corrosivity testing	8/05/08	Yes	Replacement (EU) OECD TG 431	ATLA, in press

Running numbers of ECVAM validits statements.

Noteworthy, methods 18-23 were not developed for the purpose of chemicals testing, but current validation activities explore their use for acute toxicity testing, which might lead in the short-term to an extension of the applicability domain.



Often overlooked, REACH will be applied also to new chemicals: Due to the lowered testing requirements at the typically low tonnage levels and the restrictions on any additional animal testing, new chemicals will in the future not undergo the safety evaluation that would be necessary for use as cosmetic ingredients: REACH no longer foresees any testing requirements for new chemicals produced between 100 kg and 1 ton per year and “only” two animal tests for those produced in a volume of 1 to 10 tons, i.e. the local lymph node assay (LLNA) and an acute toxicity test. Most new chemicals fall in these categories below 10 tons per year. Using the reduced LLNA (validated 2007) and the validated and accepted tiered testing strategies for acute toxicity, this means that the number of animals will be 8 mice and 8 rats on average. Noteworthy, any further testing by the chemical manufacturer would need to be suggested to the European Chemical Agency (ECHA), who after public consultation would need to respond on the testing need within two years. Thus, in practical terms, the cosmetic industry will no longer be provided by their manufacturers with sufficient chemical safety data to allow use. Thus, additional information needs require testing now falling under the scope of the 7th amendment.

For finished products, testing is rarely necessary, thus this specific ban of the 7th amendment is less critical. However, some international markets (e.g. Russia, China) require finished product testing; this incompatibility with the 7th amendment has not been fully clarified, neither for European products tested for export nor for the import of such products. The entrance to the market for American and other foreign products with new, animal-tested ingredients will be most critical, when in 2009 a marketing ban for products with ingredients tested in animals is in place. Therefore, it will be unavoidable for these foreign companies to also invest in alternatives if they want to continue to bring products to the European market. Most large multinationals are aware of this problem and started investments in alternatives more than a decade ago. However, these efforts, being company secrets, are often not available for regulatory purposes, since they represent

a competitive advantage. Therefore, it is critical that an independent validation body assures the wide applicability but also wide access to the alternative methods (Bottini et al., 2007).

While the requirements of the European market are a strong driving force for the international acceptance of alternatives, the other way around the non-acceptance of novel methods is a stumbling block for the use of alternatives, where especially chemical manufacturers do not want to carry out the alternative method for some and the traditional method for other countries. Here, the strong impact of marketing bans and/or deletion of traditional methods from test guidelines (OECD or EU) become evident. In this sense, the 7th amendment might be considered a role model for the implementation of change.

Consideration 3: Substantial progress was made towards the replacement of all toxicological tests relevant for cosmetics

Skin corrosion: The skin corrosion endpoint has been replaced, following ECVAM validation, by six validated and accepted methods (transepithelial electrical resistance test with rat skin, CorrosiTex™, four skin models). Minor follow-up work is still required not affecting the 2009 deadline. It is certainly critical that the US accepts only positive results and requires animal testing for the negative ones, exactly those substances that would typically be used in cosmetics. The planned study to address false-negative skin corrosives in the novel alternative skin irritant tests is critical to demonstrate the safety of the combined use for US acceptance. The methods primarily in use for skin corrosion require commercial, artificial human skin. To safeguard availability and free market, new products (similar methods, sometimes also referred to as “me-too” methods) require assessment of equivalence. A few more skin models are likely to be submitted as me-too methods (slight variants of validated tests, which can be assessed in a small-scale study based on performance standards) (Hartung, 2007a). Several validated methods should be available to

avoid monopolies and safeguard supply of commercial artificial human skin. This might be considered of lower priority since it is not solving a new problem, but the test developers have to be satisfied submitting their methods for validation and peer review.

Skin irritation: The endpoint has been replaced by a successful validation in 2007 with provisional acceptance by SCCP (DG SANCO’s Scientific Committee for Consumer Products, formerly SCCNFP) and inclusion in REACH test strategies. However, to enable test guideline acceptance, some small follow-up activities are required. One test was validated as a full replacement in 2007, one as a partial replacement since it was over-specific. Some additional work on an additional endpoint (IL-1 α release) to improve the assay is agreed on with L’Oréal. They have announced that they will also provide additional data on cosmetic ingredients with ECVAM quality assurance as requested for regulatory acceptance by SCCP. A test guideline has been drafted and is now submitted to the National Coordinators of the EU test guideline programme and the OECD. The validation has been submitted to IC-CVAM for expedited review. A number of me-too developments (3-4) are expected, e.g. SkinEthic RHE model, Straticell skin irritation model, CellSystems skin irritation model. This is important to maintain competition in the market and to gain acceptance. Research activities using omics technologies are being considered to reduce the false-positive rate in future test strategies. The introduction of the Globally Harmonised System (GHS) for classification and labelling might introduce an additional class of mild irritants requiring test optimisation and new validation (United Nations-GHS foresees this, but the EU-version of GHS has currently not adopted this yet).

The rabbit assay used for comparison is known to be overpredictive (60% of substances are false-positives when tested in humans). Thus, human patch test data might in the future allow an increase of the predictivity for humans. However, due to ethical constraints, it will be difficult to generate new human patch test data. With US ICCVAM, a follow-up study

was agreed to test false-negatives of the skin corrosion assay in the skin irritation assay. This shall prevent false negative skin corrosive substances from passing in a tiered *in vitro* only testing strategy.

Skin absorption: The method for this endpoint has been accepted by the OECD, actually as a political compromise in parallel to the creation of an *in vivo* method. However, no formal validation for either has taken place. It is not clear whether the 7th amendment requires such formal validation. *In vitro* percutaneous absorption data are already accepted by SCCP. Two studies have addressed the *in vitro* approaches, and the submission of results is expected shortly, financed and carried out in Germany and at EU level.

Eye irritation: This endpoint is absolutely critical for the 2009 deadline. It might in fact be considered the “holy grail” of alternative methods, because of the emotional impact for many animal lovers. With 17,000 rabbits per year (2002), animal numbers are relatively low, but the idea of a “putting chemicals into big bunny eyes” rightly creates strong emotions. Various assays are being evaluated, which, combined in a test strategy, might have the potential to replace the animal test for this endpoint. The strategy aims at utilising the strengths of specific *in vitro* assays to address required ranges of irritation and/or chemical classes (applicability domains) to classify test substances. Around 20 methods have been in industrial use for more than 2 decades and six large validation trials have evaluated them. However, no single method was found able to fully replace the complexity of the animal responses. ICCVAM and ECVAM have shared the work of collecting and compiling data for the 10 most promising methods (4 methods by ICCVAM, 6 methods by ECVAM). In addition, COLIPA is financing a statistician for ECVAM to analyse the data of the different assays described below, as well as the combination of tests based on the strategies identified in an ECVAM workshop held in 2005. Test strategy development is also supported by the Integrated Project OSIRIS (sponsored by DG RTD).

An ECVAM validation study is currently taking place for 4 cytotoxicity/-cell

function-based assays based on retrospective data. These assays seem promising for the detection of non-irritant soluble substances. The study is coordinated by an international validation management group and is based on weight-of-evidence principles (Balls et al., 2006).

The 4 complementing ICCVAM data collections for 4 other organotypic tests require reorganisation for the analysis as additional data were provided by EPAA, the European Partnership for Alternative Approaches between Commission and industry (<http://ec.europa.eu/enterprise/epaa/>). The animal data are not very reliable (Weil and Scala, 1971; see Fig. 2, page 119 in Spielmann, 1996) – actually, some people argue that if the animal experiment was better, we would have replaced it long ago. Who can reproduce something with an alternative, if the animal test itself is not reproducible? Perhaps we should have included the animal test in the earlier validation studies to show, how (badly) it performs against the historic data. An interesting option is a latent class analysis (i.e. validation without gold standard, a method used in the field of diagnostics), which might be applied here, since so many complementing methods and large datasets are available. Four methods for severe eye irritants are accepted by the EU, two of them were validated in 2007. Test guidelines have been drafted by ICCVAM for the OECD and by ECVAM for the EU. For the two validated methods, specific substance panels to test and systematically expand the applicability domain shall be carried out. A major reason for the failure of previous studies was the low quality of the animal data to compare the *in vitro* data with. A Ph.D. thesis by Susanne Scheiwiller analysed the animal test, which shall now be complemented by statistical analysis of animal data from the New Chemicals Database.

The two most promising methods based on reconstituted human corneal epithelial cells are starting validation with the support of COLIPA. These models appear promising for the identification of non-irritant neat substances like the *in vivo* application with a large applicability domain.

A submission of a refinement variant of the animal test (LVET – low volume eye test) evaluated using human data is

under peer-review by ESAC and might represent a better point of reference. Two submissions, the slug test and IRRITECTION are currently under preparation based on prospective and retrospective data, respectively, for the evaluation of completeness and adequateness of information to proceed towards peer-review.

Acute toxicity: This test could probably be considered obsolete: In the pharmaceutical area, efforts are ongoing to abandon the test, because other preliminary single dose tolerability studies and/or dose range finding data are sufficient (led by the U.K. National Centre for the 3R, NC3R). This would open a window until 2013 for cosmetics ingredients (deadline for the repeated-dose test) for cosmetics, but this deadline is likely to be postponed anyway, since no replacement is foreseeable. An analysis of the existing and new chemical database to support such a statement by ESAC was recently submitted (Prieto et al., 2008). It demonstrates the potential use of NOAEL (no observed adverse effect level) data from repeated dose to extrapolate the LD₅₀.

A validation study between ICCVAM and ECVAM to predict acute toxicity by cytotoxicity assays was completed in 2005. The study confirmed previous findings suggesting that the *in vitro* tests could identify non-toxic chemicals. These make up 80% of the new chemicals (New Chemical Database). These also make up about 70% of the old chemicals (IUCLID). A follow-up study was commissioned in 2007 to test about 60 additional chemicals to test this hypothesis. Furthermore, the test was established on the robotised testing facility at the JRC (collaboration between the Nanotechnology and Molecular Imaging unit and ECVAM), where additional substances are now being tested. COLIPA was invited to provide relevant cosmetic ingredients to complement this testing.

The Integrated Project A-Cute-Tox was initiated following an ECVAM workshop (Gennari et al., 2004). The project aims to develop and prevalidate testing strategies and will end in 2009. Two already ECVAM validated assays (CFU-GM, pyrogen tests) were shown within the A-Cute-Tox project so far to predict acute toxicity (>85%) much better than the ear-



lier cytotoxicity tests. Additional testing of these assays is planned.

The concept of toxicological thresholds of concern (TTC) (Kroes et al., 2004; Kroes, 2006) offers calculations to estimate whether a maximal exposure can result in toxicologically relevant levels in the organism. It might also be possible to demonstrate that the results of acute toxicity are not meaningful for humans: In fact, a preliminary comparison with poison centre data (publication in preparation) demonstrates that the rat is less predictive for toxic human blood levels than the cell-based tests.

Genotoxicity: Several accepted *in vitro* tests exist, but they often identify substances false-positively, which then need to be sorted out by animal tests. To omit this control would leave industry with too few options with regard to the availability of chemicals, making this more an economical than a scientific problem. A promising approach is based on an ECVAM validated method, i.e. the micronucleus test allows including this endpoint in the repeated-dose test (2013 deadline).

Various *in vitro* assays are available; however, the problem is the low quality of these non-validated tests, especially their over-sensitivity. When typically three such tests are combined, only 3% of the positive results are correct, 97% are negative in the subsequent animal test (Kirkland et al., 2005). An ECVAM workshop on the reduction of false-positives in *in vitro* genotoxicity tests was carried out (Kirkland et al., 2007) and the recommendations are being followed up with funding from COLIPA, NC3R and ECVAM. Validation of skin models for genotoxicity testing is co-sponsored by COLIPA and ECVAM.

The *in vitro* micronucleus test (MNT) was validated in 2006. This is a remarkable pilote case for various reasons: It was the first completely retrospective validation ever, i.e. as suggested by the Modular Approach to validation (Hartung et al., 2004) it made use of a compilation of existing data from various sources instead of designing a new validation study. In fact, the validation was completed in roughly half the time of a prospective study and with minimal

costs. The MNT validation also represents an example of the new close collaboration with OECD, since test guideline development was carried out in parallel to validation. Thus, consensus at OECD level already could be reached in 2008 – perhaps it would have worked out even faster if the US had been on board from the beginning: The MNT validation was one of the rare examples of validations not done in collaboration with ICCVAM, and it took about one year including an international workshop to sort out the concerns, then at OECD level, after the validity statement. This shows how much time and energy can be saved when collaborating from the start. The MNT also has a very special history with regard to acceptance: It was declared validated in late November 2006 and was entered into the final REACH legislation by the European Parliament only two weeks later. It is thus the first “legislative acceptance” in contrast to “regulatory acceptance”, since no consultation of regulators took place. Something similar is currently happening with the draft of the new EU test guideline legislation, where efforts to include the recently validated new methods for skin irritation, skin sensitization, eye corrosion and acute ecotoxicity by European Parliament are ongoing. The MNT has also already been integrated in the draft revised ICH guidelines for pharmaceuticals. Noteworthy, it is, however, only an *in vitro* test replacing another *in vitro* test of lower quality.

JaCVAM is leading the validation of the COMET assay with EU and US involvement. Both the MNT and the COMET assay promise to be more predictive. More important, however, they can be incorporated into the repeated-dose animal studies, thus making a further mutagenicity test unnecessary. The deadline for the repeated-dose toxicity test is in 2013. Therefore, the integration of the mutagenicity tests (deadline 2009) into the repeated-dose toxicity test is of high priority and would allow overcoming the use of additional animals for mutagenicity/genotoxicity testing. Study designs to validate the inclusion are currently under discussion.

ECVAM is evaluating various possibilities to reduce the number of animals in *in vivo* genotoxicity testing, e.g. the use

of only one gender, the omission of positive controls, etc. New DNA repair based assays were submitted or identified and are currently being discussed. Peptide-binding assays as entering validation for sensitization are being suggested to identify reactive chemistry, a prerequisite for mutagenicity. Combined analysis of existing mutagenicity assays instead of a battery approach has been shown (P&G) to improve their predictive capacity; the study design is under discussion.

The question might, however, be raised whether mutagenicity in human cells should be ruled out at all by an animal test. A genotoxic effect *in vitro* shows that the substance has a property which could be hazardous. Differences in the *in vivo* test can be either species-specific (rat *versus* human) or due to kinetics (does not reach the tissue at sufficiently high concentrations). These do not necessarily rule out a hazard toward humans, especially in chronic situations or hypersensitive individuals. This means that the animal experiment may possibly hide a hazard for humans.

Photogenotoxicity: This rarely tested endpoint could not be properly addressed yet, but promising approaches are being followed by COLIPA and others.

Phototoxicity: The endpoint has been replaced by a validated method for which minor follow-up activities are required. No problem for the 2009 deadline. However, the validated method is considered over-sensitive. A workshop has been suggested in which to reanalyse the test performance. This might be a first example of a revision of a validated test after some years of use, as suggested by the workshop on post-validation (Bottini et al., 2008). Work using skin models instead or together with cell cultures in a tiered testing strategy are on the way by companies and ZEBET.

In summary, what seemed a “mission impossible” when the legislation was written is almost within our grasp just six years later. It has required a multi-faceted programme (Hartung et al., 2003; Zuang and Hartung, 2005; Hartung, 2007a; Zuang et al., 2008) and close collaboration between the stakeholders. Certainly,

the *in vitro* methods have their shortcomings (Hartung, 2007b) but the same holds true for the traditional animal tests (Hartung, 2008). Sure, in some fields the harvest has not yet been brought in, but there is reason to assume that with one to two years of delay the necessary alternatives will be validated. The deadlines might even have been met if support from all sides had been instant and continuous. This is regrettable, since the cosmetic industry, which became a major driver of alternatives, would have deserved such success ready to enter into the discussion of the 2013 deadline, for which the situation is clearly less favourable.

Consideration 4: The 2013 deadline is scientifically not feasible

Regarding the 2013 deadline, there is reason for optimism in the area of skin sensitization. However, endpoints such as chronic toxicity, reproductive toxicity, toxicokinetics and carcinogenicity will require activities of a new dimension. This means that merely continuing to develop a battery of tests for organ toxicities will not suffice; instead novel approaches (systems biology, pathways of toxicology, etc.) need to be included in a collaboration between the stakeholders (Hartung and Leist, 2008; Leist et al., 2008). Still, it is necessary to seriously attempt a solution of this problem. Motto: Failure is an option, not trying it is not! A vision outlined recently by the US National Academy of Science (National Research Council, 2007) can only be realised with an effort of the dimension of the Human Genome Project, but discussions toward such an approach are taking place on both sides of the Atlantic. It is clear that, at present, we are faced with basic scientific shortcomings that prohibit complete replacement of the endpoints. However, advances in science, technologies and toxicological approaches are opening up new possibilities to tackle these fields. It is clear that full replacement of these endpoints will not be achieved in due time (i.e. 2013). However, at present, targeted large initiatives aiming at bringing all stakeholders together in large European research efforts (DG RTD Innovative

Medicine Initiative and some Integrated Projects like Carcinogenomics <http://www.carcinogenomics.eu/>, the Assuring Safety without Animal Testing (ASAT) initiative <http://www.asat-initiative.eu/>, COLIPA etc.) are assessing the application of current scientific knowledge, new technologies, new cell systems and new endpoints (focussing on critical processes) for these areas.

It is highly unlikely that it will be possible to predict chronic toxicity with any test strategy or battery of non-animal tests. For reproductive toxicity, some possibilities might emerge from the Re-ProTect project (2004–2009, <http://www.reprotect.eu/>). Cancer bioassays are very unlikely to be requested for cosmetic ingredients, since chemicals identified as positive in mutagenicity/genotoxicity assays are usually abandoned. Notable exceptions are hair dyes and some antimicrobials, which, owing to the reactive chemistry required for their use, are often positive in mutagenicity assays and have raised cancer concerns. However, in case the carcinogenic potential needs to be evaluated, cell transformation assays, which are currently under validation, might be used. Promising alternative methods exist for skin sensitization (tests currently under validation might allow the identification of large groups of non-sensitisers), and the Integrated Project Sens-it-iv seeks new methods (2005–2010, <http://www.sens-it-iv.eu/>).

Skin sensitization: Skin sensitization is a systemic endpoint and requires at least two expositions to a substance, owing to which it belongs to the category of repeat-dose toxicity and falls under the 2013 deadline. The Local Lymph Node Assay, an ICCVAM validated refinement method, was scientifically endorsed by the ECVAM Scientific Advisory Committee (ESAC) and is now an OECD adopted Test Guideline (TG 429). Non-radioactive variants are currently under development. ECVAM has received the submission of one of these variants for evaluation and it is foreseen that other submissions will follow. For this reason an ECVAM workshop was organised (Basketter et al., 2008) to discuss how to define criteria for their evaluation with the test method developers and experts in

the area. However, these represent refinement methods for REACH, not solutions for cosmetics after 2013.

A number of relevant workshops have been organised (Casati et al., 2005; Kimber et al., 2007; Basketter et al., 2007; Gerberick et al., 2008) to identify the most promising methods and to give advice on how to progress towards their validation (peptide-binding assay, skin permeability, dendritic cell based methods). Three promising methods, the peptide-binding assay, the U937 test and the hCLAT test are currently under development/optimisation within COLIPA; ECVAM is playing an advisory role in this phase. Recently, an evaluation of the status of development of these methods was performed, and it is likely that they will be able to enter a formal validation study within this year.

The Integrated Project Sens-it-iv was set up (<http://www.sens-it-iv.eu/>) following an ECVAM workshop (Casati et al., 2005) with results expected by 2010. Four different cell systems and four different markers have shown promising results. These are going to be standardised and assessed in an inter-laboratory trial. Beside Sens-it-iv, the general interest in developing relevant alternative methods for skin sensitization testing is supported by the large investment by industry in this area. COLIPA has just published a call for research proposals that will increase the mechanistic understanding of how chemical allergen exposure impacts upon cutaneous cell types and/or lead to the development of cell-based assays for the prediction of chemical-induced skin sensitization (budget 2.5 million €).

Toxicokinetics: Toxicokinetics is best defined as an adjunct technique to support the interpretation of toxicological results in risk assessment. Two areas must be considered, i.e. developing and refining the quantitative aspects of *in vitro* testing, and facilitating and promoting the use of toxicokinetic prediction techniques in risk assessment. In line with the workshop on this theme, held in 2007 (Bouvier d'Yvoire et al., 2007), there is a need for a more precise estimation of *in vitro* bioactive concentrations. The so-called nominal concentrations, obtained by simply dividing the amount



of compound added by the volume of the test system, can be misleading for compounds that bind strongly to components of the test system, in particular proteins or lipids. A more precise characterisation of the test systems and the use of modelling techniques to estimate the active concentrations *in vitro* are necessary. This is because the conditions in which the potentially toxic compounds act can be very different between the *in vitro* and *in vivo* situations. The development of a formal approach is crucial in this respect and has the potential to increase the confidence in quantitative *in vitro* toxicology results. It is also a prerequisite for the acceptance of quantitative *in vitro* to *in vivo* extrapolation for risk assessment purposes, i.e. beyond the stage of hazard identification.

At present, animal-free techniques of prediction of pharmacokinetics or toxicokinetics are used as a screen in pharmaceutical research. They rely essentially upon the use of physiologically based pharmacokinetic (PBPK) modelling techniques, with input of quantitative parameter values generated by various techniques *in silico*, *in vitro*, *ex vivo* and *in vivo*. This field is still under development, and the use and acceptance of PBPK in risk assessment is increasing. However, the field is still a long way from full replacement. Large scale research is necessary to reach the level of confidence and acceptance of these techniques demanded by an objective of full replacement, and it is difficult to predict now when or even whether this objective can be achieved at all. The important research need in the field is exemplified by the large projects dedicated to kinetics and toxicology prediction in the Framework Programmes and related initiatives like the Technology Platform Innovative Medicine Initiative (<http://www.imi-europe.org/>).

Carcinogenicity: ECVAM is carrying out the validation of three cell transformation assays parallel to OECD test guideline development. Parallel work by JaCVAM on a fourth variant with ECVAM in the advisory board is ongoing. The tests have the potential to partially replace the 2-year bioassay in the rat. With up to 1 million € per substance, the animal assay is so costly that it is hard-

ly carried out (from almost 5,000 new chemicals notified in Europe over the last 25 years, 14 had a cancer bioassay in the new chemical database). Thus, any *in vitro* test would first of all enable a larger number of substances to be tested at all. However, the general concerns about the high false-positive rate of the cancer bioassay (some aspects summarised in Hoffmann and Hartung, 2006) prompt a need for new approaches. New tests based on the loss of gap-junctions have been proposed, for which pre-validation needs to be initiated.

Reproductive toxicity: Due to the complexity of reproductive toxicity and the variety of toxicological mechanisms involved, in the medium term a (partial) replacement can only be achieved by focussing on toxicological targets with high prevalence. This requires setting priority on the relevant target cells and/or biological mechanisms, since the complete mammalian reproductive cycle cannot currently be mimicked *in vitro*. In order not to compromise consumer safety, the use of *in vitro* tests for assessing reproductive toxicity of cosmetic ingredients therefore depends on a detailed analysis of historical *in vivo* data of reproductive toxicants belonging to the various chemical groups relevant for cosmetics. In this context, the histopathological data of reproductive organs in repeated dose experiments are of high interest, since this endpoint seems to be very sensitive. A comprehensive database should provide information on which cells/tissues and which cell functions need to be assessed *in vitro* depending on the various chemical classes.

Within FP 6 the Integrated Project “ReProTect” has been established in order to develop a toolbox of *in vitro* tests assessing various cell functions relevant for mammalian reproduction, such as hormone production, germ cell maturation, uterine function, etc. However, since not all possible target cells/mechanisms are covered within this project, further research activity is necessary. For successfully developed tests from ReProTect, the relevance and reliability of the test has to be assessed for cosmetic ingredients in (pre)-validation studies.

Repeated-dose toxicity: The Integrated Project Predict-IV (<http://www.predict-iv.toxi.uni-wuerzburg.de/en/>) addresses strategies to improve the assessment of drug safety in the early stage of development and in the late discovery phase, by an intelligent combination of non animal-based test systems, cell biology, mechanistic toxicology and *in silico* modelling, in a rapid and cost effective manner. Since the assessment does not differ for cosmetic ingredients, the project is fully relevant also for this sector. The project will integrate new developments to improve and optimise cell culture models for toxicity testing and to characterise the dynamics and kinetics of cellular responses to toxic effects *in vitro*. It was based on two ECVAM workshops on chronic toxicity and physiology-based pharmacokinetic modelling (PBPK) (Prieto et al., 2006; Bouvier d’Yvoire et al., 2007).

The FP7 Predict-IV project will address one of the major problems of *in vitro* methods, the absence of data for extrapolations: from *in vitro* studies, NOAELs, serving as starting points for extrapolations, cannot be easily derived, since systemic doses cannot be directly compared to concentrations of a substance applied to cells. In cell systems which best represent *in vivo* target organs, the most predictive endpoints indicative of adverse effects will be used to determine the no-observed-effect-concentrations (NOEC) *in vitro*. The NOEC will be based on the estimation of measured *in vitro* intracellular concentrations of the drug and/or its relevant metabolites. These NOECs will then be transformed to doses received using appropriate modelling techniques, in particular advanced PBPK modelling including Monte-Carlo techniques. Since the model systems will be based on human cells and the PBPK-models incorporate human parameters and potential interindividual differences in humans, the need for extrapolations will be reduced and NOAELs can be predicted.

Repeated-dose systemic toxicity requires a new systematic approach in order to move from a traditional, animal-based assessment to a mechanism-based approach. A vision was created last year by the US National Academy of Science (National Research Council, 2007). In

essence, it is suggested to move from animal to human cell based approaches using pathways of toxicology. Emerging technologies are systems biology (a bioinformatics guided combination of several “omic” approaches), high-throughput and high-content testing as well as computational toxicology (Hartung and Leist, 2008; Leist et al., 2008). The US has advanced this concept in its ToxCast program (<http://www.epa.gov/comptox/toxcast/>). Notably, proposals have been furthered toward the US Congress to initiate a project similar to the Human Genome Project, where the effect of small molecules (either intended, i.e. pharmacology, or unintended, i.e. toxicology) on gene expression and function shall be studied. Such a project of several hundred million € set up as an international collaboration, might bring a new dimension to our understanding of the interaction of man with his chemical environment.

The US approach that is currently shaping can be characterised as “bottom-up”, i.e. it is based on data generation from thousands of substances in a broad array of systems allowing later data-mining. Due to its nature, this approach is far from regulatory applications and test development. It is suggested that this “bottom-up” approach be complemented by a “top-down” approach, i.e. an approach guided by far fewer testing chemicals with a clear quality-controlled toxicological profile in test systems, which have been shown to be relevant as predictors of human toxicity. It will be important to safeguard quality assurance (Good Cell Culture Practice, (Coecke et al., 2005)), validity and regulatory usefulness. Furthermore, with a view to the regulatory implementation, the principles of evidence-based toxicology (Hoffmann and Hartung, 2006; www.ebtox.org) should be implemented from the start.

Ecotoxicology: The cosmetics regulation does not cover environmental issues, but chemical ingredients need to be tested for ecotoxicity, creating needs for animal testing of cosmetic ingredients and making them relevant for cosmetics. Activities in ecotoxicology/environmental toxicity are currently focused on short-term (acute) aquatic toxicity and bioconcentration. Endpoints such as

long-term (chronic) aquatic toxicity testing or toxicity to birds have not yet been tackled. Changes in legislation, e.g. environmental assessment of pharmaceuticals and REACH, have increased and will increase the numbers of fish used. However, regarding cosmetics and the use of chemicals tested under REACH, it should be borne in mind that only substances produced/imported in volumes >10 t are tested in fish.

The *Threshold Approach* (= Upper Threshold Concentration [UTC] step-down approach, (Jeram et al., 2005)) for acute aquatic toxicity testing is based on a retrospective analysis of data in the New Chemicals Database carried out by ECVAM and the ECB. Its validity for the reduction of the numbers of fish was endorsed by ESAC in 2006. It is part of the intelligent testing strategy in the requirements for REACH. The Commission has submitted a proposal to the OECD to incorporate this strategy into the OECD Testing Guideline Programme. The success of this project in the OECD heavily relies on swift recruitment of one member state competent authority to carry out this work. Otherwise the European Commission will have to retire the project, which would save about 200,000 fish and 20 million € of testing costs for REACH!

The most promising replacement method is the Fish Embryo Test (FET), for which a lot of data are available and which is already used for effluent testing in Germany. Since cosmetic ingredients are tested as chemicals and might be produced above volumes of 10 t, the FET is interesting as a replacement; noteworthy, P&G and L'Oréal are putting a lot of effort into the proper development of this test. In 2006, Germany submitted a background review document on this test to the OECD and to ECVAM. The OECD has established an *ad hoc Expert Group for the Fish Embryo Test (FET) Test Guideline*. First analysis of the dataset showed that gaps have to be filled before the validity of the FET can be evaluated. Several studies are in the planning phase and ECVAM was asked to give scientific advice.

A study using fish cell lines was finalised this year showing that fish gill cell lines might be useful for acute aquatic toxicity testing, e.g. in a testing strat-

egy. In addition, ECVAM is part of the advisory board of a CEFIC-LRI project investigating the use of fish cell lines and fish embryos.

The endpoint bioconcentration assesses whether a substance accumulates in fish. In a first step, bioconcentration models based on physico-chemical properties and physiological parameters (uptake, metabolism and excretion) are used. If the outcome is above a certain threshold, a test in fish becomes mandatory. The current strategy followed by industry and also proposed in the intelligent testing strategy for data requirements under REACH is to improve the existing *in silico* methods by including *in vitro* data on metabolism. Substances which are metabolised do not accumulate in the organism.

Consideration 5: Regulatory acceptance is (once again) the bottle-neck

For cosmetics, first of all the regulations for chemicals have to be applied. Relevant testing procedures are accepted by the EU (so far Annex V of the Dangerous Substance Directive via the National Coordinators of the Test Guideline Program, in the future by ECHA and the EU Test Guideline Regulation via appropriate procedures) and the OECD. Specifically for cosmetics, the DG SANCO Scientific Committee for Consumer Products (SCCP) has a regulatory function, but normally relies on the test guidelines of EU/OECD. However, the chemicals regulation is about hazard, whereas the cosmetics evaluation is about risk assessment; this last part is covered partially by the SCCP with regard to the ingredients regulated in the Annexes of the Cosmetics Directive. Incorporation into the Annexes of the European chemicals legislation and OECD test guidelines may last longer than the validation process. By direct collaboration, the process has been smoothed on both the European and the OECD level. Measures taken over the last five years include:

- Collaboration with the National Coordinators of the EU Test Guideline Program, which for example led to the acceptance of positive results from 4



tests for severe eye irritants in 2004 by the Member State Competent Authorities supported by an ECVAM survey; the omission of test guidelines from REACH made the ongoing creation of an EU test guideline regulation necessary; it has to be assumed that the panel will be maintained also because of its parallel role toward the OECD Test Guideline Program.

- Collaboration with ICCVAM, the US counterpart, which is critical for the global applicability and acceptance of alternatives: mutual representation on the scientific advisory committees, various joint workshops and studies. This collaboration is one of four pilot projects in the EU/US cooperation in the implementation of the Guidelines for Regulatory Cooperation and Transparency agreed under the Transatlantic Economic Partnership of 1998.
- Collaboration with OECD: secondments of three ECVAM staff members, current validations parallel to three test guideline developments by OECD (carcinogenicity, mutagenicity, endocrine disrupters), observer status of OECD on the ECVAM scientific advisory board since 2005. ECVAM largely contributed to an OECD guidance document on Good Laboratory Practice (GLP) for *in vitro* studies accepted in 2004 (OECD, 2004) and to guidance document on validation in 2005 (OECD, 2005).
- As a major breakthrough, OECD has finally accepted two ECVAM validated full replacements for animal tests (skin corrosion and phototoxicity).
- The vice-chair of SCCP (DG SANCO's Scientific Committee for Consumer Products) is member of ESAC, OECD, ICCVAM and JaCVAM are observers. Two presentations by ECVAM to SCCP and a meeting with the joint chairmen of the DG SANCO scientific committees in May 2007 increased information exchange.
- In September 2007, the International Collaboration on Cosmetic Regulation (ICCR) was created between US, Canada, Japan and EU. Noteworthy, harmonisation of alternative methods represents a key topic from start. ICCR encouraged exploring the opportunity to create an International Council of Validation Bodies (ICVaBo) to synergise further.

However, despite all progress, regulatory acceptance represents one of the bottle-necks and a pacemaker for the availability of alternatives. A close collaboration with ex-ECB/EChA and EU/OECD test guideline development assuring wide-applied use of these guidelines is crucial. The collaboration with SCCP might be intensified.

Regulatory acceptance of ECVAM validated tests relevant for cosmetics: OECD/EU test guidelines: 3 tests acute toxicity, 6 tests skin corrosion, 1 test skin sensitization, 1 test phototoxicity; under discussion: 1 strategy for acute aquatic toxicity, 1 strategy skin sensitization, 2 tests skin irritation, 1 test mutagenicity (already accepted in REACH), 2 tests eye corrosion. However, several OECD accepted methods are not replacements (e.g. acute toxicity OECD guidelines are *in vivo* tests), which means they can be used only until the relevant deadline, and ecotoxicology is not covered by the 7th amendment.

This means that of 35 ECVAM-validated methods 11 are accepted for chemicals/cosmetics (notably 7 are accepted for pharmaceuticals/vaccines and 13 are in the process of regulatory acceptance, most expected for 2008). Noteworthy, the test strategies for REACH developed in RIP 3.3 foresee many of the validated methods as well as some under validation. Published by the agency, this can be considered also as regulatory acceptance. Recently, some methods originally developed for drugs and biologicals (CFU-GM test, pyrogenicity tests) turned out to have potential for acute toxicity testing of chemicals and cosmetics.

Regulatory acceptance is often the bottle-neck for the use of alternative approaches. These issues have been addressed in a workshop on post-validation (Bottini et al., 2008) and by EPAA, but they deserve further attention. The international dimension (see above and Bottini et al., 2007) makes this a complicated issue, but this also means that much progress can be made here.

Consideration 6: The collaboration between the EU Commission and the cosmetic industry has paved the public/ private partnership EPAA

Given an average of three years, three laboratories and 300,000 €/method of funding for a full prospective validation study, it became instantly evident that this cannot be supported for the large number of tests to be validated by JRC/ECVAM alone. Thus already in December 2002, ECVAM initiated the creation of an industry partnership ("MILAN – More Input Less Animal Network" – proposal taken up by Unilever and others in 2003). The creation of EPAA was a proposal I first made in a meeting between DG ENTR and the JRC in April 2005, when meeting in preparation for the first "Europe goes alternative" conference. DG ENTR then invited twelve companies short-listed by ECVAM to a meeting with Vice-President Verheugen in August 2005, which resulted in a steering group that elaborated the 3R declaration and the structure for EPAA. The goal is to create a sustainable political momentum by starting a partnership between the European Commission (DGs ENTR, JRC, RTD, ENV and SANCO) and industry with an action program and annual review mechanism of the progress on alternative, non-animal testing approaches. This was accomplished by a written commitment (signed declaration) by industry to work intensively towards the development, validation and implementation of alternative testing methods and an action program to implement it within a short, medium and longer term.

The Partnership was officially launched on 7 November 2005 at the "Europe goes alternative" conference (http://ec.europa.eu/enterprise/events/animal_tests/index_en.htm) by Commissioners Verheugen and Potocnik and industry representatives. Its purpose is to promote the development of new "3R" methods (refine, reduce, replace) as modern alternative approaches to safety testing. The European Commission will ensure the secretariat of the Partnership. However, with regard to funding, EPAA remained below (my personal) expectations so far, in contrast to the substantial

investment by COLIPA's Steering Committee on Alternatives to Animal Testing SCAAT program at the same time. The work of EPAA is still very much restricted to workshops, the annual conference and consensus documents. No substantial funding of practical work has taken place and efforts to make data available for validation are very limited. Industry is willing to be active in the implementation of the action plan, but all industry representatives insisted on the need for the Commission to take the lead (coordinating role, mainly by the DG ENTR run secretariat) of this project as it involves different sectors (chemicals, pesticides, pharmaceuticals, cosmetics, medical devices, food and biotechnology) and past experience has proven their difficulties to coordinate themselves.

The EPAA Partnership brings together different industrial sectors, allowing a holistic and more innovative approach to safety testing using the best available science and expertise. This Partnership encourages (<http://ec.europa.eu/enterprise/epaa/>):

- The promotion of industry activities and investments in 3R research.
- A more rationalised implementation of regulatory testing requirements.
- A more streamlined process for the acceptance of scientifically validated alternative testing approaches.
- The identification of needs for research in alternative safety testing methods and facilitation of relevant multi-stakeholder research projects.
- The sharing of knowledge and best practice between sectors in implementing the 3Rs Declaration agreed at the "Europe Goes Alternative" conference in 2005.
- The consistent communication on research and implementation of the 3Rs in relation to safety assessment.

The EPAA Action Programme is designed around five main themes, with more sub-activities:

- Mapping and evaluating past and current 3R activities.
- Prioritising and implementing research based on the 3Rs.
- Best practice implementation.
- Implementation of the 3Rs in Regulation and Decision-Making.

- Validation and acceptance of new and alternative test methods and strategies.

The Action Programme, comprising short, medium and long-term activities, will be reviewed and updated every year. Implementation will be ensured through Working Groups in which stakeholders will be involved.

Companies and stakeholders can participate by:

- Providing expertise,
- Joining research programmes,
- Becoming involved in pilot programmes,
- Providing project and/or financial support, where necessary and possible.

I think it is fair to say that the impulse to create EPAA was first taken up by the cosmetic industry. However, the remarkable buy-in from all industrial sectors and diverse Commission services and continuing work shows how this has spilled over to others. The foundation of EPAA was clearly laid in the close collaboration between ECVAM and the DG ENTR unit/directorate responsible for cosmetics legislation and the expert stakeholder taskforce to establish the timelines for the 7th amendment. Thus, the 7th amendment represents the starting point for this important private/public partnership. Another reason to be happy about the political "victim"...

Consideration 7: How to measure success so far with regard to 7th amendment?

Table 3 summarises the main activities in the field. I have indicated in colours the expected timing relative to the deadlines of the 7th amendment and also given a very personal appraisal for the probability that this will lead to a scientifically acceptable replacement of the animal test. To quote a popular saying (which is attributed both to Nils Bohr and to baseball coach Yogi Berra, <http://letterfromhere.blogspot.com/2006/12/bohr-leads-berra-but-yogi-closing-gap.html>): "It's tough to make predictions, especially about the future". Thus we should focus on where we are now.

The extensive elaboration of provisional time-lines for phasing out animal experiments for the Cosmetics Directive by DG ENTR and ECVAM published by the Commission in 2004 suggested that the prerequisites for the first deadline in 2009 could largely be met given optimal support and instant regulatory implementation. However, these conditions were not always met, and some replacements will now only be available after 2009. The deadline for 2009 can now no longer be achieved for eye irritation, genotoxicity as well as photogenotoxicity, and probably acute toxicity. There are, however, prospects for 2010-2011. The "softer" deadline (i.e. which can be further postponed in a co-decision procedure) in 2013 represents an enormous challenge, which is addressed e.g. in the Integrated Projects initiated.

The question arises, how to measure the success so far? Some views on this:

Full replacement? Full replacement achieved on the basis of ECVAM validation for the 2009 deadline (6 out of 8 toxicological endpoints in time, 2 slightly delayed with some risks for failure):

In time: skin corrosion, skin penetration, phototoxicity, genotoxicity, skin irritation and possibly acute toxicity

Delayed: Eye irritation, photogenotoxicity and probably acute toxicity

For the 2013 deadline there are prospects for 1 (sensitization) out of 5, but relevant efforts are ongoing in all areas.

Timing? The 2009 deadline will be largely met, the missing endpoints might be completed by 2010-2011, but regulatory acceptance might take more time.

Animal reduction?

Skin corrosion: 100% (3 animals to 0).

Skin irritation: 100% (3 animals to 0).

Eye irritation: 5% for sorting out severe irritants (still 1-3 animals used).

Acute toxicity: 83% (45 animals to 8 on average, possible abandonment by 2009).

Genotoxicity: 100% (screening, but animals used for confirmation of positives).

Skin sensitization: 60% (20 animals to 8, note: 2013 deadline).

Photogenotoxicity/phototoxicity: no standardised animal test (saving unclear).

In consequence, the classic test battery ("six-pack", i.e. skin corrosion, skin

**Tab. 3: Summary of status of individual toxicological endpoints**

Colour coding: Over-all appraisal (Thomas Hartung) on the endpoint in first row, whether full replacement can be reached (in time: green, somewhat after the deadline: yellow, unlikely: red, white for reduction and refinement approaches); with regard to timing for each approach (discrepancies: e.g. although for cancer some results are expected in time, the likelihood to achieve with these full replacements is low), not colour-coded for reduction and refinement approaches, which are, however, relevant until 2013 or as long as deadline is postponed.

Toxicological endpoint	Approach	Replacement	Likelihood of meeting deadlines (personal estimate)	Timing, Limitations	
Skin corrosion	Skin models	Full	Done	2004 (OECD)	
	Skin irritation	Full	Done	2007 (ESAC), regulatory acceptance pending, minor follow-up work	
Skin absorption	Human and animal skin	Full	Done	2004 (OECD)	
	Skin models	Full	50%	2008-9 (ESAC)	
Eye irritation	Organotypic models (4)	Full as part of strategies	50%	2007 (ESAC) for severe, 2009-2010 for mild	
	Cell-based models (4)	Full as part of strategies	30%	2009-2010	
	Human corneal epithelial models (2)	Full as part of strategies	80%	2010	
	Slug test, Irritation test	Full as part of strategies	30%	2010	
	Low volume eye test	Refinement	50%	2008 (only until 2009 and as point of reference)	
Acute toxicity	Tiered testing strategies	Reduction (45 → 8)	Done	2002 (OECD), 2007 (ESAC) until 2009 only	
	Abandon when repeated dose available	Full	80%	2008	
	Cytotoxicity for non-toxic substances	Full for non-toxic (70-80%)	50%	2009	
	Functional assays (2)	Full	50%	2009-2010	
Test strategies from A-Cute-Tox		Full	80%	2010-2011	
	Genotoxicity	Various existing tests	Full (animal still used to reduce false-positive rate)	Done	Before 2000 (false-positive rate to be reduced)
		Micronucleus test and COMET assay (integration into repeated dose)	Integration into a test falling under 2013 but likely to be postponed	80%	2009-2010, basic MNT 2006 (ESAC, REACH)
Reduction approaches		No	50%	2008 (only until 2009)	
Repair-based assays (1-2)	Full	30%	2010-2011		
Photogenotoxicity	Pilot studies	Full	50%	2010-2011, no standardised animal test	
Phototoxicity	Fibroblasts	Full	Done	2004 (OECD), false-positive rate	
	Skin models	Full	80%	2009-2010	
Skin sensitization (2013)	Local Lymph Node Assay (LLNA)	Refinement & reduction (20 → 16)	Done	2002 (OECD), animal test	
	Reduced LLNA	Reduction (16 → 8)	Done	2007 (ESAC), until 2013 only	
	Non-radioactive LLNA	Refinement	80%	2010	
	Peptide-binding assay	Full as part of strategies	50%	2010	
	Dendritic cell tests (2)	Full as part of strategies	30%	2010	
	Strategies from Sens-it-i.v.	Full	30%	2011-2012	
Toxicokinetics (2013)	Various uptake, metabolism and barrier models, PBPK models	Partial	10%	Unlikely, no standardised animal test	
Carcinogenicity (2013)	Cell transformation assays (2)	Full as part of strategies	80%	2013	
	Cell contact assay	Full as part of strategies	10%	2012-2014	
	Strategies from Carcinogenomics	Full	10%	2012-2015	
Reproductive Toxicity (2013)	Extended one generation study	Reduction (3,200 → 1,200)	80%	2009-2010, until 2013 only	
	Endocrine disruptors (12)	Full as part of strategies	80%	2009-2011	
	Embryotoxicity (3)	Partial	Done	2002 (ESAC), does not fully cover the animal test	
	Strategies from ReProTect	Full	<10%	Unlikely	
Repeated dose toxicity (2013)	Strategies from Predict-IV	Full	<10%	Unlikely	

irritation, eye irritation, acute toxicity, genotoxicity, and skin sensitization) was reduced from about 73 to 18 animals.

Animal suffering? Two major contributions are the refinement of the skin sensitization test (lymph node swelling instead of skin reaction) and for eye irritation the pre-screening for severe eye irritants with the tests validated in 2007.

Costs? The alternative methods validated for cosmetics are largely cost-neutral regarding the assays for the 2009 deadline. However, there are possibly enormous savings by methods and test strategies relevant for the 2013 deadline, given that some of the animal tests cost 50,000 to 700,000 € per substance.

Number of ESAC statements (cumulative) relevant for cosmetics?

1997:	1
1998:	4
2000:	6
2006:	9
2007:	17
2008:	18
2008-2010:	up to 68

(Expected from running validations: 3 skin irritation, 1 refinement eye irritation, 3 carcinogenicity, 2 skin absorption, 2 reduction mutagenicity, 6 acute toxicity, 12 eye irritation, 12 endocrine disrupter, 3 sensitization refinement, 3 sensitization, 1 phototoxicity, 2 mutagenicity).

It needs to be noted, however, that on the one hand the number of endpoints replaced is more important than the number of tests validated, i.e. to have several tests for the same endpoint does not add a new quality. On the other hand, often methods complement each other (practical availability, applicability domain, components of a testing strategy) and help to avoid monopolies in case of commercial methods.

In conclusion, the 7th amendment has prompted the largest efforts ever to make validated alternative methods available. It is remarkable, that an industry with a relative small animal use in collaboration with the Commission has driven a development, from which larger animal users such as chemical industry for REACH but also other sectors benefit. Acknowledging all uncertainties for the outcome of the initiated programme, it still appears that the 7th amendment has shaped

the way safety assessments are done and which role alternative approaches play.

References

- Balls, M., Amcoff, P., Bremer, S. et al. (2006). The Principles of Weight of Evidence Validation of Test Methods and Testing Strategies. The Report and Recommendations of ECVAM Workshop 58. *ATLA* 34, 603-620.
- Basketter, D., Pease, C., Kasting, G. et al. (2007). Skin sensitization and epidermal disposition: The relevance of epidermal disposition for sensitization hazard identification and risk assessment. The report and recommendations of ECVAM workshop 59. *ATLA* 35, 137-154.
- Basketter, D., Cockshott, A., Corsini, E. et al. (2008). An evaluation of performance standards and non-radioactive endpoints for the local lymph node assay. The report and recommendations of ECVAM workshop 65. *ATLA* 36, 243-257.
- Bottini, A. A., Amcoff, P. and Hartung, T. (2007). Food for thought... on globalization of alternative methods. *ALTEX* 24, 255-261.
- Bottini, A. A., Alepee, N., De Silva, O. et al. (2008). Optimization of the post-validation process. The report and recommendations of ECVAM workshop 66. *ATLA* 36, 353-366.
- Bouvier d'Yvoire, M., Prieto, P., Blaauboer, B. J. et al. (2007). Physiologically-based Kinetic Modelling (PBK Modelling): Meeting the 3Rs Agenda. The Report and Recommendations of ECVAM Workshop 63. *ATLA* 35, 661-671.
- Casati, S., Aeby, P., Basketter, D. A. et al. (2005). Dendritic cells as a tool for the predictive identification of skin sensitization hazard. *ATLA* 33, 47-62.
- Coecke, S., Balls, M., Bowe, G. et al. (2005). Guidance on Good Cell Culture Practice. *ATLA* 33, 261-287.
- Eskes, C. and Zuang, V. (2005). Alternative (non-animal) methods for cosmetic testing: current status and future prospects. *ATLA* 33 *Suppl. 1*, 1-228.
- Gennari, A., Berghe, C., Casati, S. et al. (2004). Strategies to Replace In Vivo Acute Systemic Toxicity Testing. *ATLA* 32, 437-459.
- Gerberick, F., Aleksic, M., Basketter, D. et al. (2008). Chemical Reactivity Measurement and the Predictive Identification of Skin Sensitisers – The Report and Recommendations of ECVAM Workshop 64. *ATLA* 36, 215-242.
- Hartung, T., Bremer, S., Casati, S. et al. (2003). ECVAM's response to the changing political environment for alternatives: Consequences of the European Union chemicals and cosmetics policy. *ATLA* 31, 473-481.
- Hartung, T., Bremer, S., Casati, S. et al. (2004). A Modular Approach to the ECVAM Principles on Test Validity. *ATLA* 32, 467-472.
- Hartung, T. (2007a). Food for thought ... on validation. *ALTEX* 24, 67-72.
- Hartung, T. (2007b). Food for thought ... on cell culture. *ALTEX* 24, 143-147.
- Hartung, T. (2008). Food for thought ... on animal tests. *ALTEX* 25, 3-9.
- Hartung, T. and Leist, M. (2008). Food for thought ... on the evolution of toxicology and phasing out of animal testing. *ALTEX* 25, 91-96.
- Hoffmann, S. and Hartung, T. (2006). Towards an evidence-based toxicology. *Human Exp. Toxicol.* 25, 497-513.
- Jeram, S., Riego Sintes, J. M., Halder, M. et al. (2005). A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notified in the European Union. *Reg. Tox. Pharmacol.* 42, 218-224.
- Kimber, I., Agius, R., Basketter, D. et al. (2007). Chemical respiratory allergy: Opportunities for hazard identification and characterisation. The report and recommendations of ECVAM workshop 60. *ATLA* 35, 243-265.
- Kirkland, D., Aardema, M., Henderson, L. and Müller, L. (2005). Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. I. Sensitivity, specificity and relative predictivity. *Mutation Res.* 584, 1-256.
- Kirkland, D., Pfuhler, S., Tweats, D. et al. (2007). How to reduce false positive results when undertaking in vitro genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop. *Mutation Res.* 628, 31-55.
- Kroes, R., Renwick, A. G., Cheeseman, M. et al. (2004). Structure-based



- thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food Chem. Toxicol.* **42**, 65-83.
- Kroes, R. (2006). The threshold of toxicological concern concept in risk assessment. *Toxicology Lett.* **164**, S48.
- Leist, M., Hartung, T. and Nicotera, P. (2008). The dawning of a new age of toxicology. *ALTEX* **25**, 103-114.
- National Research Council, Committee on Toxicity and Assessment of Environmental Agents. (2007). Toxicity Testing in the Twenty-first Century: A Vision and a Strategy, <http://www.nap.edu/catalog/11970.html>
- OECD (2004). Series on principles of good laboratory practice and compliance monitoring, Number 14, Advisory Document of the Working Group on Good Laboratory Practice, The Application of the Principles of GLP to in vitro Studies, [http://appli.oecd.org/olis/2004doc.nsf/linkto/env-jm-mono\(2004\)26](http://appli.oecd.org/olis/2004doc.nsf/linkto/env-jm-mono(2004)26)
- OECD (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Environmental Health and Safety Monograph Series on Testing and Assessment No. 34. Available: http://www.oecd.org/document/30/0,3343,en_2649_34377_1916638_1_1_1_1.00.html
- Prieto, P., Baird, A. W., Blaauboer, B. J. et al. (2006). The assessment of repeated dose toxicity in vitro: a proposed approach. *ATLA* **34**, 315-341.
- Prieto, P., Bulgheroni, A., Kinsner-Ovaskainen, A. et al. (2008). Estimation of acute oral toxicity using the 28-days repeated dose studies in rats (submitted).
- Spielmann, H. (1996). Alternativen in der Toxikologie. In F. P. Gruber and H. Spielmann (eds.). *Alternativen zu Tierexperimenten*. 108-126 Verlag.
- Weil, C. S., and Scala, R. A. (1971). Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. *Toxicol. Appl. Pharmacol.* **19**, 276-360.
- Worth, A. P. and Balls, M. (2002). Alternative (non-animal) methods for chemical testing: current status and future prospects. *ATLA* **30 Suppl. 1**, 1-125.
- Zuang, V. and Hartung, T. (2005). Making validated alternatives available – the strategies and work of the European Centre for the Validation of Alternative Methods (ECVAM). *AATEX* **11**, 15-26.
- Zuang, V., Eskes, C., Griesinger, C., and Hartung, T. (2008). ECVAM key area topical toxicity: update on activities. *AATEX* **14**, S523-S528.

Acknowledgement

More than ever I have to acknowledge the work of ECVAM staff and collaborators. I have used some unpublished reports prepared by among others Annamaria Antonella Bottini, Michel Bouvier d'Yvoire, Susanne Bremer, Silvia Casati, Sandra Coecke, Raffaella Corvi, Chantra Eskes, Laura Gribaldo, Claudius Griesinger, Marlies Halder, Sebastian Hoffmann, Annett Janusch-Roi, Agnieszka Kinsner, Jens Linge, Cristian Pelizzer, Costanza Rovida, and Valerie Zuang. I owe them not only for this contribution but also for the continuous discussion, but I take personal responsibility for the content and how it is presented.

Correspondence to

Prof. Dr. Dr. Thomas Hartung
 EU Joint Research Centre
 Institute for the Protection and Security of the Citizen (IPSC)
 Traceability, Risk and Vulnerability Assessment (TRiVA) unit
 21027 Ispra
 Italy
 e-mail: Thomas.Hartung@ec.europa.eu